The Dissertation Committee for Michael B. Cammarata Certifies that this is the approved version of the following dissertation:

Development of Top-Down Mass Spectrometry Methods for Structural Characterization of Protein Macromolecules Utilizing 193nm Ultraviolet Photodissociation

Committee:

Jennifer S. Brodbelt, Supervisor

Richard M. Crooks

Lauren J. Webb

Andrew D. Ellington

Walter L. Fast

Development of Top-Down Mass Spectrometry Methods for Structural Characterization of Protein Macromolecules Utilizing 193nm Ultraviolet Photodissociation

by

Michael B. Cammarata, B.S. , M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2016

This dissertation is dedicated to my family, Glenda Boyer, Don Cammarata, and Matthew Cammarata.

Acknowledgements

First, I would like to thank my PhD advisor Jennifer Brodbelt for allowing me to be an independent scientist and take scientific risks throughout my course of research. Without her guidance and constant support it would not have been possible to bring my ideas to light. I would also like to thank my collaborators which provided great scientific problems and even better samples for me to work with. Specifically, thank you to my collaborators Chris Schardon, Dr. Walter Fast, Dr. Ross Thyer, and Dr. Andrew Ellington here at UT and those at University of Connecticut.

I also want to thank the Brodbelt Group and in particular Dr. Buddy O'Brien who started doing native-MS work in the lab. Additionally, I would like thank Dr. Lindsay Morrison who discussed my research problems with me to find better solutions.

Finally, I need to thank my friends and climbing partners, Hank Legan, Hannah Crews, and Max LaBerge, for their friendship and healthy distractions over the years. Last but not least, my partner Brittany Laurent for her incredible support and unceasing thoughtfulness that has helped push me forward to accomplish my goals.

Development of Top-Down Mass Spectrometry Methods for Structural Characterization of Protein Macromolecules Utilizing 193nm Ultraviolet Photodissociation

Michael B. Cammarata, PhD

The University of Texas at Austin, 2016

Supervisor: Jennifer S. Brodbelt

The dissertation will discuss the advancement of informative structural biology techniques utilizing a top-down centric workflow with 193nm ultraviolet photodissociation (UVPD) mass spectrometry. Native electrospray ionization is used to transport proteins to the gas phase in a native-like state, then UVPD is used for structural characterization to reveal ligand binding sites within a protein-ligand complex as well as detect conformational changes based upon the suppression or enhancement of backbone cleavages. Conformational changes induced by ligand exchange or removal and single amino acid mutations as well as combinations of the two (ligands and mutations) are investigated. The rich fragmentation patterns of UVPD are also used for structural characterization of crosslinked proteins. Typically these crosslinking experiments are performed by bottom-up mass spectrometry with has significant shortcomings. The main drawback is the need for proteolysis which cuts proteins into small peptides, thus increasing the complexity of the samples and its subsequent analysis. Additionally this proteolysis step loses the post-translation modification information or amino acid mutations that may be driving a specific protein-protein interaction. Top-down methods

avoid protein digestion and thus are used to directly evaluate the protein interactions or protein complexes. These two methodologies will bring the mass spectrometry and structural biology community a step closer to the realization of high-throughput structural biology for proteins and their interactions with other proteins and small molecules.

Table of Contents

Chapter 1: Introduction

1.1 Introduction

Deciphering the structure of macromolecules such as proteins and DNA has consistently remained one of the most harrowing problems to be solved in the world of biochemistry. This dissertation focuses on developing mass spectrometry methods for protein structure characterization. To have a full appreciation for development of these new methods it is imperative that the history of using mass spectrometry in the context of structural biology needs to be assessed. There are four main categories of mass spectrometry methods applied for structural biology: covalent labeling for surface accessibility, covalent crosslinking of proteins, hydrogen-deuterium exchange, and native mass spectrometry. Two of these, covalent crosslinking and native mass spectrometry, were explored and improved upon with a top-down centric focus in the present dissertation.

With the advent of x-ray crystallography[1] the dream of generating three-dimensional structures from precise analytical data had been achieved and since then has been adopted widely as the gold standard across the field of structural biology field as the gold standard.[2,3] Additionally, nuclear magnetic resonance (NMR) has played a key role in determining in solution structures of proteins.[4,5] While both of these methods have the potential to yield single amino acid resolution and side chain position, they also require potentially prohibitively large amounts of protein, up to mM concentrations and/or mg quantities, for data acquisition. While these techniques yield very high-resolution structures, they do not always reflect the native structure of the protein found at biologically relevant concentrations.[6,7] These issues can lead to protein precipitation for

NMR and packing artifacts at key loops in the protein structure during crystallization. In addition to these issues, data analysis is complex and time-consuming, a factor that keeps these techniques from achieving a higher-throughput that could be used for screening the interactions of proteins containing unique amino acid substitutions with specific drugs or ligands or other proteins. This is the intersection of structural biology where mass spectrometry can excel. Owing to its sensitivity, use of sub-microgram amounts of proteins, and its comparative ease of data analysis, mass spectrometry has many attributes that showcase its performance for protein structural analysis. The ability of mass spectrometry to quickly screen many combinations of protein-ligand interactions and to evaluate changes as a function of protein sequence or ligand identity compensates for the fact that the data obtained from mass spectrometry methods is low resolution and does not translate to three-dimensional representations. With clever new strategies that combine labelling methods, computational modelling, and mass spectrometry detection , many low to mid resolution structural biology techniques have been developed over the past two decades. A number of these established and emerging methods will be surveyed in the following sections.

1.2 Covalent Labeling

Covalent labeling is precisely what the name implies, using covalent labeling chemistry to decorate reactive and accessible amino acids on protein surfaces with trackable mass tags (ones that can be monitored by mass spectrometry based on mass shifts). Two generalized strategies are possible: solvent accessibility measurements[8] and chemical crosslinking.[8,9] With solvent accessibility, the protein is modified on its exposed surface by reactive small molecules. The labelled protein is typically digested into more easily detected peptides which are separated and analyzed by LCMS/MS, and the

quantities of modified and unmodified sites are monitored based on the constituent peptides.[8] The variations in the amounts of modified versus unmodified sites reflects the accessibility of different regions of the protein. On the other hand crosslinking uses fixed length crosslinking ligands to join together two available, reactive amino acids of a protein or pair of proteins in solution that are in close spatial proximity. Subsequent mass spectrometric identification of the reacted amino acids is used to develop distance constraints for molecular modeling[10] and determining protein-protein interactions for protein networking,[11] as well as determining large macromolecular structures.[12]

1.2.1 Solvent Accessibility

Most commonly surface accessibility experiments take advantage of specific side chain chemistry to selectively label amino acids.[8] There are numerous solvent accessibility probes that have been utilized for this type of mass spectrometry workflow, and several examples are shown in Figure 1.

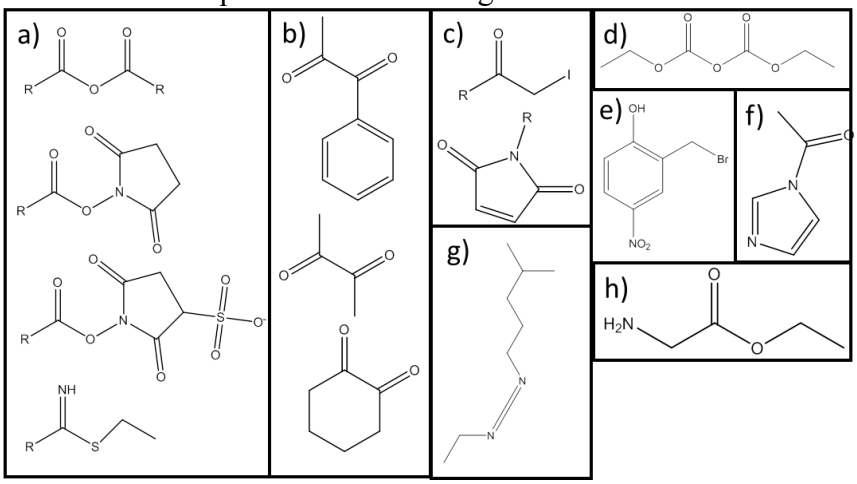

Figure 1. Common reagent(s) for covalent labeling of proteins that target (a) lysine, (b) arginine, (c) cysteine, (d) histidine, (e) tryptophan, (f) tyrosine, and (g,h) acidic residues. EDC (shown in g) activates acidic residues prior to reaction with (h).

The ones in Figure 1a target primary amines (Lys and N-terminus) which is the most common reactive site for solvent accessibility studies. The free amine on the N-terminus and side chain lysine are particularly popular targets due to their high nucleophilicity. For example, historically acetic anhydride was used as a surface accessible labeling reagent.[8,13] Another very popular way to label lysines comes from N-hydroxysuccinimidyl (NHS) leaving group chemistry which has seen widespread use for these types of experiments.[14,15] While lysine chemistry is very convenient there have also been specific chemistries developed for nearly all amino acid functionalities. These include modifications for acidic residues (glutamic and aspartic acid)[8,16] (Figure 1g,h), arginine (Figure 1b),[8,17,18] cysteine (Figure 1c),[8,19] histidine (Figure 1d),[8,20] tyrosine (Figure 1f),[8,21] and tryptophan (Figure 1e).[8,22] There have also been several other chemistries optimized to be less specific labeling reagents using standard solution chemistry like diethylpyrocarbonate, DEPC, which has been shown to react with lysine, tyrosine, cysteine, arginine, serine, and threonine.[23] Using less specific labels allows more sites to be modified in a single experiment, thus providing a greater array of solvent accessible site information at the expense of more complicated mixtures of peptides and more complex data analysis.

There are two newly emerging techniques using laser activation to label non-specifically across nearly all amino-acids which yields the most extensive amount of information as well as the fastest labeling time scale. This latter feature reduces the chance of structural distortion that slow labeling techniques may induce during the labeling process.[15,24] The two methods are laser-activated carbene footprinting[25,26] and fast photochemical oxidation for proteins (FPOP) (Figure 2).[27,28] Carbene footprinting uses a sterically strained diazirine ring or one that is photoreactive with UV light .[9,25,26] A

355 nm laser is pulsed for a fixed amount of time to activate the reagent in solution, and it reacts with accessible amino acids on a very rapid time-scale. Alternatively, the protein solution can be mixed with millimolar concentrations of peroxide which upon laser (250-270 nm) irradiation produces OH radical species (this is the FPOP method).[28–30] These OH radical species react with any solvent-exposed residue that can be oxidized. Residues Gln, Glu, Asp, Asn, Ala, and Gly are relatively unreactive to hydroxyl radicals, but virtually all other amino acids are reactive. Activated carbenes react with all amino acids and the protein backbone itself.[26]

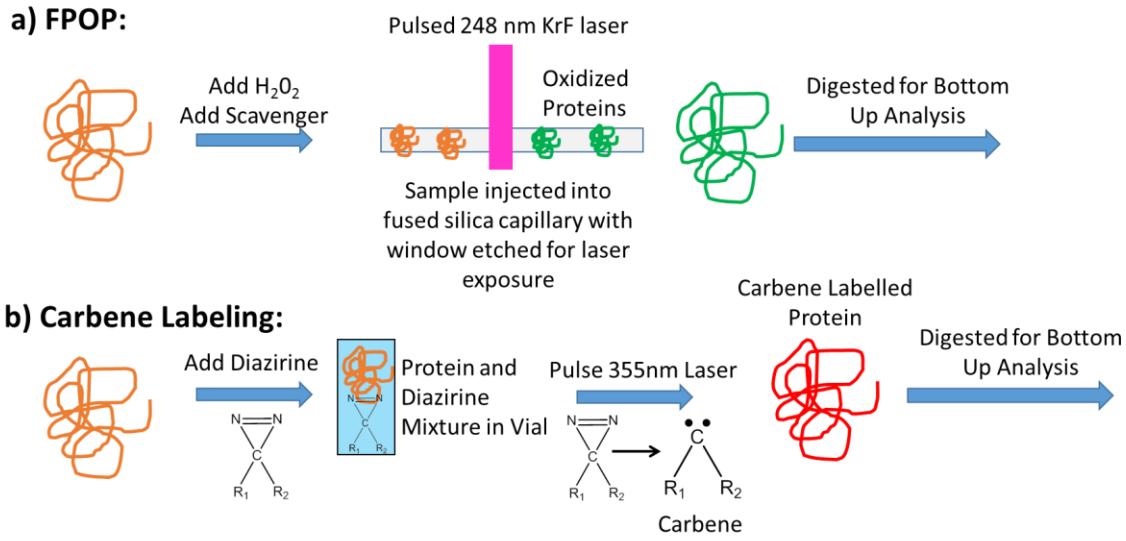

Figure 2. Generalized workflow for (a) fast photochemical oxidation of proteins (FPOP) and (b) carbene labeling.

After the protein system has been labelled using one of the methods described above, it is generally digested using a specific protease such as trypsin to produce smaller peptides which are more easily analyzed than intact proteins. The modified peptide mixture is injected, separated, and peptides are identified by LC-MS/MS.[8] Here the identified modified peptides are quantified relative to the corresponding unmodified

peptides in order to evaluate the reactivity of the various side-chains; it is this relative reactivity which correlates with solvent accessibility (Figure 3a). Based on the quality of the MS2 spectra the reactivities of different sites (which are directly related to the rates of modification of each site) can be derived from the fragmentation patterns which display mass shifts in the fragment ions depending on the location of the chemical probe. There are also other solvent accessibility techniques that involve the use of isotopically encoded reactive tags which are used under a combination of denaturing and native conditions to label the protein.[8,15,31]

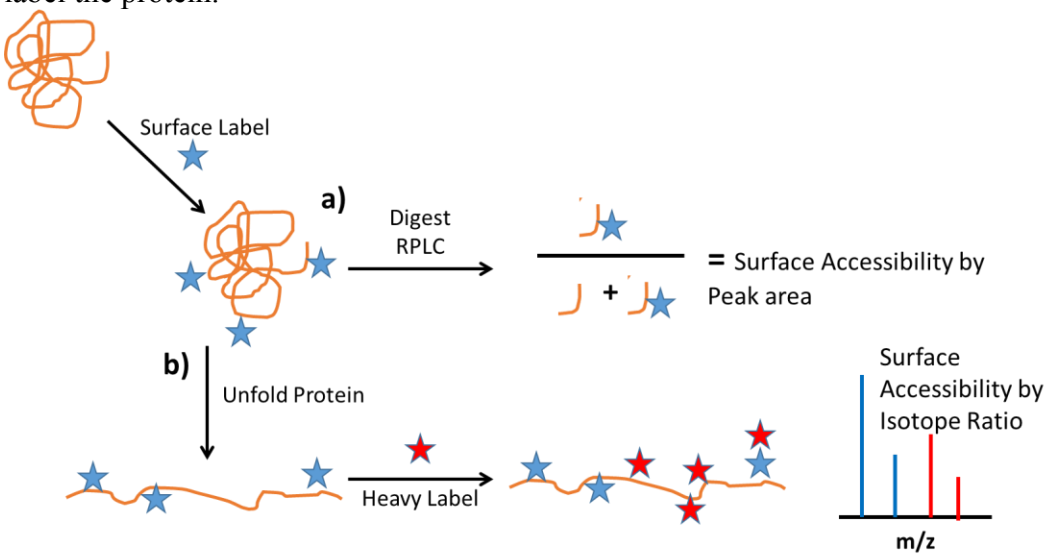

Figure 3. Generalized workflow for surface accessibility measurement using (a) peak areas and (b) isotope ratios of labelled proteins.

This generates a mixture of fully labelled peptides (denaturing conditions, one isotopic version of the reagent) versus partially labelled peptides (native conditions, a second isotopic version of the reagent), and the extent of modification in the native state is determined by comparison of the the peak abundances between the isotopically light and heavy generated versions of peptides (Figure 3b). This offers a more direct option for data processing in comparison to the methods that don't utilize isotopic labelled reagents

and thus handle quantitation based on quantitative comparison of modified and unmodified peptides.[8]

1.2.2 Crosslinking Mass Spectrometry

Alternatively to individually labeling reactive amino acids for information about their surface availability, a dual reactive covalent crosslinker can be used to determine residues near each other in two dimensional space.[9] The length of the crosslinking arm provides a distance constraint between the reacted residues. These distance constraints can be used to aid structural modeling of proteins[32] and protein complexes[12] as well as provide a way to determine protein interaction networks when used in conjunction with more elaborate pulldown techniques used to isolate arrays of interacting proteins.[11,33,34]

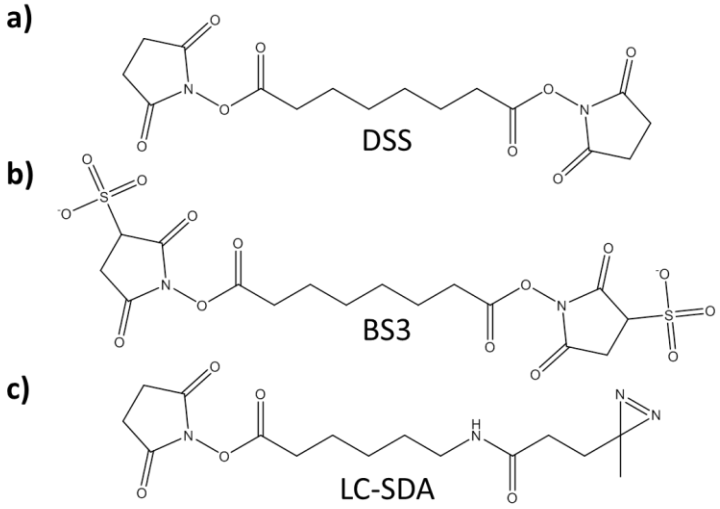

Figure 4. Examples of hombifunctional crosslinkers (a) disuccinimidyl-suberate (DSS) (b) bis(sulfosuccinimidyl) suberate (BS3) and (c) heterobifunctional crosslinker succinimidyl 6-(4,4'-azipentanamido)hexanoate) (LC-SDA).

As with solvent accessibility measurements these chemical crosslinkers come in a wide assortment of flavors that can be customized for the the structural goal. The most commonly used crosslinker is disuccinimidyl suberate (DSS) which takes advantage of

NHS leaving groups to join two primary amine groups together.[35] Bissulfosuccinimidyl suberate (BS3) is a water soluble analog of disuccinimidyl suberate (DSS) (Figure 4). These linkers both have a total distance constraint of 30 Å from Cα to Cα of the reactive residues (Figure 5).

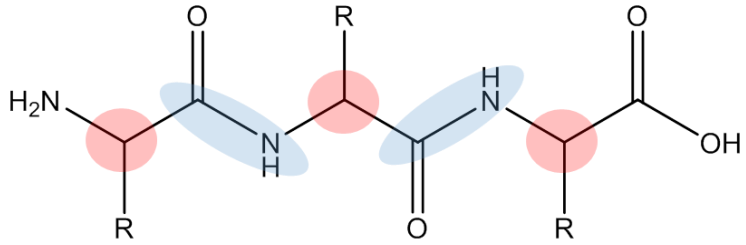

Figure 5. A generalized tri-peptide. The blue ovals indicate the peptide bonds and the red spheres indicate the alpha-carbons.

The distance constraint is estimated by adding the length of the fully extended crosslinker to the length of the amino acid sidechains plus adding approximately 3Å for each residue to factor in protein dynamics.[36] While BS3 and DSS are the simplest examples of homobifunctional crosslinkers, there are a myriad of heterobifunctional crosslinkers which can be useful to cover different distances or to target other side-chains.[9] One other option to broaden the information obtained from crosslinking strategies is to use photoactive crosslinkers which are often significantly more promiscuous. These crosslinkers contain aryl azides, diazirines, or benzophenones which react with numerous protein sites.[9,37,38] Typically with this strategy the crosslinker is heterobifunctional, and one reactive functionality targets a specific site to allow the crosslinking to be anchored to the protein in a characteristic manner. Then the partially modified protein is exposed to UV light to trigger the secondary reaction step. There is another class of crosslinkers which stitch together acidic residues (Asp, Glu, C-terminus) by means of EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide) coupling to form a new peptide bond (Figure

1g).[9] These crosslinks are called zero-length crosslinkers since there is no inserted length once the new peptide bond is formed. These reactions are ideal to determine spatially close residues or those that may be interacting.[9]

While crosslinking experiments can provide a large wealth of information, the crosslinks are notoriously difficult to identify due to low reaction yields, and the fragmentation patterns of the crosslinked peptides are particularly difficult to interpret. The low reaction yields are due to the fact that during crosslinking reactions three reaction pathways may occur: no reaction, dead-end resulting from only a single reactive site with the other site hydrolyzing, and the successful crosslinking reaction. Upon digestion at least two different categories of crosslinked products may be detected: intra-peptide crosslinks and inter-peptide crosslinks.[9,39] The successful crosslinking reaction occurs in low yield relative to no reaction at all or dead-end pathways. To attempt to overcome these drawbacks, tri-functional crosslinkers have been generated to allow enrichment of the crosslinks of interest (Figure 6a).[40] These are either functionalized directly with a biotin group attached with a spacer arm for streptavidin enrichment,[40] or they are functionalized in a bio-orthogonal manner (such as click-chemistry) which can followed up with biotin attachment and crosslink enrichment.[9] These strategies work well to reduce the overall complexity of the samples by removing unmodified peptides, but there still remains a mixture of dead-end and crosslinked peptides. Tri-functional crosslinkers also tend to be fairly bulky, in comparison to the bi-functional analogs, which can block access to reactive residues by steric factors.

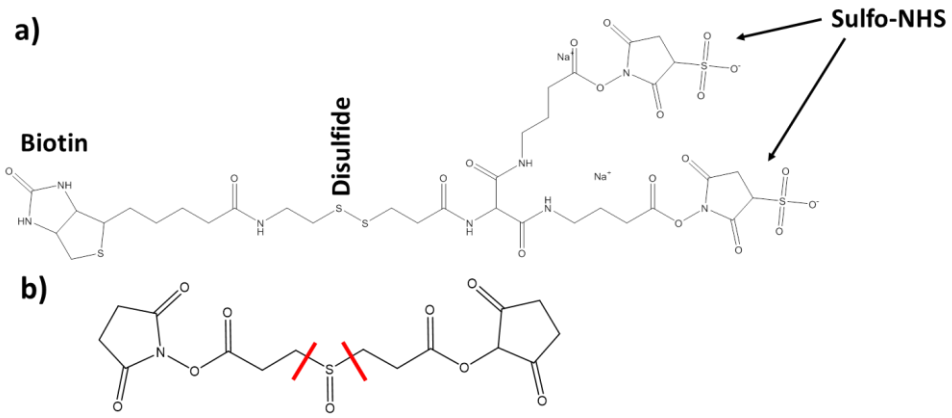

Figure 6. Designer crosslinkers with homobifunctional attachment allowing (a) biotin enrichment and release through disulfide reduction and (b) CID specific fragmentation which is signified in red lines (DSSO).

To attack the other main problem (overly complex peptide fragmentation which confounds interpretation), another type of trifunctional crosslinker has been developed. Identifying the two crosslinked peptides is challenging because the total mass of the crosslinked product is comprised of two unknown peptides from different regions of the protein(s) plus the mass of the crosslinker. Many peptide combinations might fit the mass of the crosslinked product. Even with tandem mass spectrometry, the resulting fragments are difficult to assign to one specific chain of the peptide and not the other.[39] This problem can be partially resolved with the use of a high mass accuracy mass spectrometer to allow determination of a well-defined chemical composition of the crosslinked product. However, even high accuracy assignments fall short for more complexes mixtures of proteins that are subjected to crosslinking. In cases such as these, gas-phase specific fragmentation chemistry is embedded into the crosslinker, essentially via the incorporation of a cleavable bond in the crosslinker (Figure 6b).[11,39] These crosslinkers are designed to undergo selective bond cleavage when subjected to collisional activation, thus releasing the individual peptides that contribute to the overall

mass of the initial crosslinked product.[11,39] Then MS3 can be used to sequence the constituent peptides. This strategy also has the ability to distinguish between intra-crosslinked and inter-crosslinked peptides and those with dead-end modifications. The most popular choice of gas-phase cleavable crosslinker is disuccinimidyl sulfoxide (DSSO) which fragments specifically at the neighboring bonds of the sulfoxide group.[39]

1.2.3 Hydrogen-deuterium Exchange

Hydrogen-deuterium exchange (HDX) is another form of covalent labeling; however, due to the nature of the experiment it is a completely reversible reaction.[8,41] This technique was born out of the classic HDX experiments performed with NMR.[42] In HDX experiments proteins are exposed to high levels of deuterium oxide in solution. The amide hydrogen atoms of the protein backbone exchange as well as the side-chain hydrogen atoms on respective heteroatoms (N,S,O) with the deuterium atoms of D_2O.[41,43] The extent of deuteration per backbone position reveals the amount of protection or accessibility of the residue and its availability for exchange. The protected residues are either hydrogen-bonded to either the backbone or side chain or buried within the core of the protein. Those residues that appear unprotected (highly exchangeable) are in disordered regions or not involved in hydrogen bonds.[44] HDX consistently gives the highest resolution structural information among all mass spectrometry labelling methods.[43,45]

While this technique is superior to covalent labeling due to the high resolution structural information, there are some serious technical obstacles that must be overcome owing to the prevalence of back exchange once the protein is labelled with deuterium.[41,43] Side-chain exchanges are typically not observed in solution due to very rapid back-exchange in solution upon contact with non-deuterated solvents, with the exception of

histidine.[46] Briefly, the general workflow for a HDX experiment involves labeling the intact protein or proteins with deuterium oxide for a fixed amount of time followed by an acid quench and rapid freeze to discourage back exchange.[41,45] The resulting mixture is then digested quickly by pepsin, a non-specific protease that is active at 0°C at pH 2. These attributes of pepsin are advantageous to obtained high percentages of overlapping sequence coverage and allow proteolysis to be performed at conditions that minimize back-exchange. The peptide mixture is separated over a reversed phase column with a fast gradient (typically under 15 minutes) to reduce peptide back exchange during chromatography.[41,47] From the MS1 spectra collected, the extent of deuteration can be calculated by the isotopic mass shift of each peptide towards higher masses.[44,45] MS2 spectra are generally collected for peptide identification but not per residue deuterium content due to hydrogen scrambling by collisional activation methods.[48,49] Interestingly, it has been recently shown that electron-based activation (such as electron transfer dissociation) of deuterium exchanged peptides and proteins does not cause an appreciable amount of hydrogen scrambling during fragmentation.[50,51] The general HDX workflow described above, which was once a large technical challenge, has been commercialized by Waters Corporation and LEAP Technologies, allowing widespread adoption of this technique. The commercial platforms offer automated sample handling including automation of exchange and the quenching, with additional automation for on-line pepsin digestion followed by injection onto the column that is cryogenically cooled to reduce back exchange during separation.

HDX experiments are commonly collected over multiple time points for the HDX reaction; by using a broad range of HDX reaction times, a kinetic map of deuterium uptake into the protein backbone can be derived.[43–45] HDX can also be applied in a pulsed type mode with a fixed time of labeling over the course of time of a reaction or during

folding/unfolding to probe specific protein conformational changes on a temporal dimension.[45,52] An application of this is the study of conformational changes due to amyloid aggregation in solution over time, during which the amyloid aggregate size can be monitored by turbidity of the solution or dynamic light scattering experiments.[52,53] HDX has also been adopted heavily by the pharmaceutical industry for epitope mapping between antibody and antigen as well as comparing biosimilar drugs.[45,54] Briefly, this is accomplished by measuring the exchange rates of the separate antibody and antigen and then monitoring the exchange rates upon their complexation. Those peptides that display a decrease in exchange rates are indicative of being involved in the epitope interaction.

1.3 Native Mass Spectrometry

Native mass spectrometry is the study of biological molecules sprayed and ionized from salt buffers in an attempt to mimic physiological solution conditions.[55,56] Unlike all the other mass spectrometry techniques discussed above, this technique does not involve protein digestion or labeling. Native mass spectrometry is a modified top-down centric mass spectrometry technique in which intact proteins are analyzed by MS1 and MS2 methods in an effort to derive stoichiometric and conformational insight. The goal of top-down mass spectrometry is to examine intact proteins without the need to reconstruct whole proteins from peptide pieces.[57] Native mass spectrometry has the ability to build directly upon the advantages of top-down mass spectrometry to study intact proteins, protein-ligand complexes, and protein-protein complexes.[55,56]

Changing solution conditions dramatically alters the resulting charge states of the protein. In denaturing solvents, proteins are observed in a large range of high charge states , but proteins in native solutions generate a narrow range of low charge states (Figure 7).[58] Typically denaturing solutions consist of >50% organic solvent with a low

percentage (<2%) acid component and no salts. Native solution are typically close to 100% aqueous with a relative high concentration of a volatile salt: ammonium acetate, ammonium bicarbonate, or ammonium formate in concentrations ranging from 5 mM to 2 M depending on the ionic strength required to retain the protein interactions.

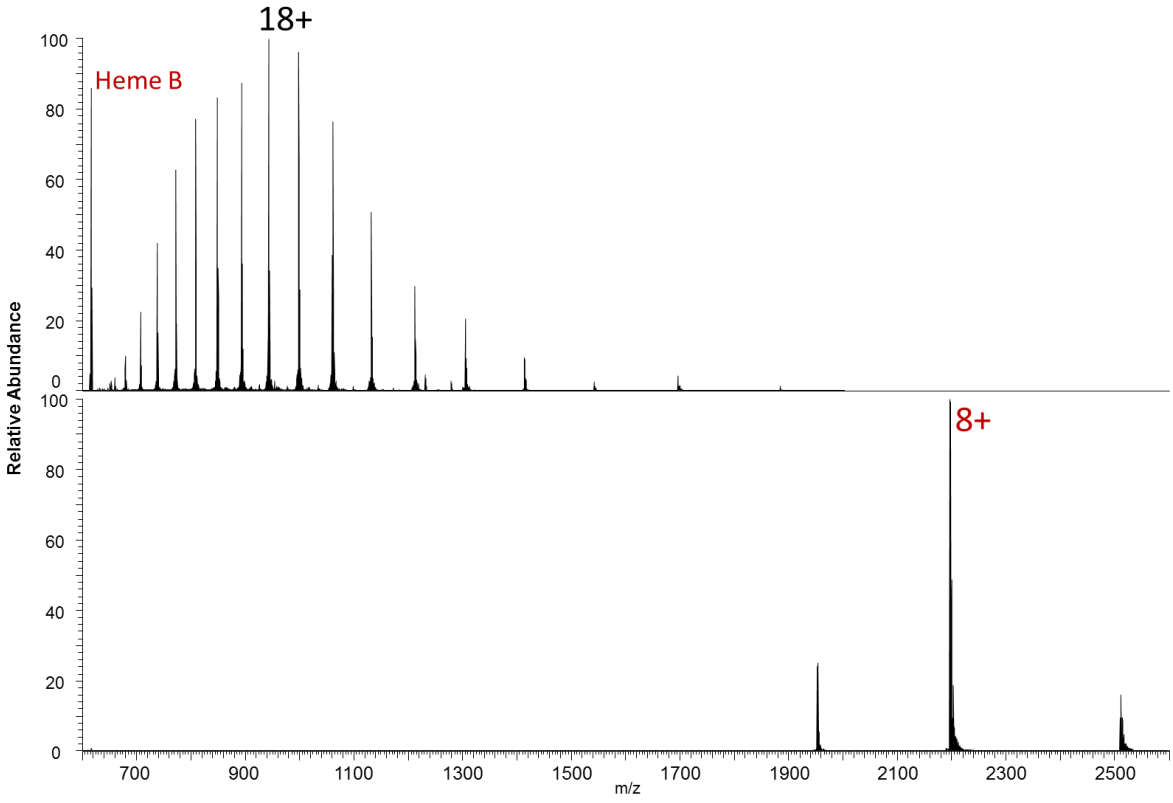

Figure 7. Myoglobin sprayed under denaturing conditions in (a) 50% water:50% acetonitrile:0.1% formic acid. It can also be seen sprayed under native conditions in (b) 50 mM ammonium acetate at neutral pH. The heme group in retained during the native spray.

These solvent changes are thought to alter the mechanism of ionization; protein ions from denaturing solutions follow the chain ejection model (CEM) whereas proteins from native solution experience the proposed charged residue model (CRM).[58] The CEM suggests that the denatured protein unfolds, and the hydrophobic residues, which were

found in the interior region of the protein in the native state, arrange to the outer edge of the charged, hydrophobic droplets formed by the ESI process.[58] Upon droplet shrinking owing to multiple coloumbic explosions, the hydrophobic portion of the protein will extrude itself through the barrier of the droplet and pick up surface charges resulting in the formation of highly charged protein ions. For the alternative CRM, the protein folds are retained and the protein remains near the center of the droplet.[58] Upon continuous evaporation of the droplet, the charge from the surface of the droplet is deposited onto the surface of the protein, leading to a low number of charges on the protein.[58] The narrow charge state distribution also provides support for the retention of a natively folded protein. While the mechanism of native spray and how native the gas phase structure is a hotly contested subject overall[58–61], there is a growing amount of evidence supporting the premise that native-like protein structures are retained during the native spray process.[62,63] The other evidence includes retention of non-covalent interactions upon ionization[34,64,65] as well as collisional cross section measurements of the native protein charge states that align well with theoretical collisional cross sections based on solution models.[66–68]

One of the common applications of native mass spectrometry is to monitor the stoichiometry and topology of large protein complexes.[69] These range from the size of a hexamer comprised of two hetero-trimers (toyocamycin nitrile hydratase, 85.6 kDa) [70] to ATP-synthase (538 kDa) [71] to matured virus capsids (3-4 MDa).[63] One general method for analysis of these large macromolecular complexes is use time of flight (TOF) mass spectrometry to measure the masses of the intact complexes.[55,56,69] The resulting complexes can be mass selected and subjected to collisional activation to cause disassembly of the complex.[69,71] This dissociates the complex into smaller subunit complexes, some which retain protein-protein interactions, or ejection of monomer proteins from the complexes.[69,71] Those smaller complexes generated by collisional

activation are indicative of protein-protein interactions within the larger complex. This information can be used to reconstruct the overall topology of the protein complex. A new alternative technique uses high energy collisions with a surface, surface induced dissociation (SID).[72,73] With this technique the selected protein complex is accelerated into a large surface which offers fast activation and high energy deposition.[73] SID has been shown to selectively break the weakest interface of the complex without unfolding the protein, an outcome that correlates with the total surface area of the protein interfaces and ultimately gives information about subunit structure (architecture).[74] In contrast, conventional collisional activation of protein complexes proceeds by unfolding of a monomer subunit followed by ejection from the complex.[70,74,75] An alternative approach to examining protein topologies by native mass spectrometry involves deliberate step-wise disruption of the proteins complexes by addition of organic solvents to the solutions.[76] In this experiment a gradient of organic solvent is added to the protein solution and the MS1 is repeatedly recorded. This method disassembles the complexes to reveal the topology of the protein complex.[76]Topologies solved via this solution disruption method versus those cleaved by MS/MS methods in the gas phase have been shown to agree well.[74,76]

Native mass spectrometry can be used to monitor protein-ligand interactions as well,[62,64,65,77] a task that is very useful for screening protein-drug interactions[78] and to optimize the use of lipids for enhancing the analysis of membrane proteins.[62,79,80] Some labs have been successful in using native mass spectrometry to measure dissociation constants of challenging protein systems.[81–83] Very recently new methods have been developed to measure the thermodynamic binding constants of a series of membrane protein – lipid complexes.[77] The study of protein-ligand interactions has also lead to understanding of biological mechanisms. For example, the mechanism of an ABC

transporter (TmrAB) that was unclear by crystallographic methods was proven by native mass spectrometry experiments.[80] It was determined that this ABC transporter needed dual ATP binding and hydrolysis to transport lipid-A type molecules across the membrane.[80]

Another avenue for studying protein-ligand interactions by native mass spectrometry involves the use of high energy and/or electron activation (as opposed to low energy collisional activation) techniques to localize the binding sites.[75] High energy or electron activation of a protein ligand complex can cause covalent bond cleavage while retaining non-covalent protein-ligand interactions that are confirmed based on observation of fragment ions that retain the ligand even when the backbone of the protein has been cleaved.[64,65,84–88] These ligand-retaining fragments are referred to as holo fragments. Three specific high energy or electron-based activation techniques that exhibit this phenomenon are electron-capture dissociation (ECD), its non FT-ICR analog electron transfer dissociation (ETD), and ultraviolet photodissociation (UVPD). Briefly, ECD uses high energy electrons from a heated filament to interact with the ion of interest, such as proteins.[89,90] This activation method yields mainly *c/z* type fragments ions. ETD entails the transfer of an electron from a fluoranthene anion to the positively charges ion, also resulting in product ion of c/z ions.[91] In the UVPD process, an ion may absorb photons from a laser, ultimately resulting in fast, high energy activation that generates *a/b/c* and the *x/y/z* type ions with unprecedented sequence coverage for top-down analysis.[92,93] With each of these techniques yielding holo fragments, it is possible to reconstruct the site of non-covalent interactions between the protein and the ligand of interest.[64,85,87,94] UVPD in particular is the subject of much of the research in the present dissertation.

In addition to using these activation techniques in conjunction with native mass spectrometry to localize ligand binding, it has also been shown that these fast high energy activation techniques can be used to study conformational changes of proteins.[94–96] Conformational changes of proteins can originate from ligand binding, amino acid substitutions, or changing initial solvent conditions. Additionally, ETD has been recently shown to be a valuable tool to determine salt bridges in the gas phase.[97]

1.4 Overview of Chapters

Despite the versatility and impressive capabilities of the various mass spectrometry methods summarized above for the characterization of biological macromolecules, there remains room for many more advances in the technologies and applications. This dissertation will focus on the development and applications of nativespray mass spectrometry for an array of structural biology problems. Chapter 3 will demonstrate the adaptation of crosslinking to a top-down type platform in contrast to the commonly used bottom-up type platform. The impetus for this work is to characterize complexes via a top-down approach to offer a more powerful means to pinpoint crosslinks. This top-down approach should also be well-suited for examination of crosslinks in proteins that contain post-translational modifications, ones that might be missed using bottom-up methods for which complete sequence coverage is rarely attained.

The remainder of the dissertation focuses on using native-MS UVPD methodology to probe protein-ligand complexes, particularly to characterize the locations of bound ligands (co-factors or inhibitors) and to map conformational changes that arise from ligand binding. The complex loop rearrangements of dihhydrofolate reductase (DHFR) during its catalytic cycle, which results in the conversion of dihydrofolate to

tetrahydrofolate, have been studied extensively using conventional methods such as NMR and X-ray crystallography.[98,99] Thus, DHFR serves as an excellent benchmark system to explore the use of UVPD to monitor conformational changes upon binding of the co-factor (NADPH), the substrate (DHF) and an inhibitor (MTX), as described in Chapter 4. The capabilities of UVPD-MS to characterize conformational changes as a function of point mutations and the impact of inhibitor binding is expanded for DHFR and two clinically isolated mutants in Chapter 5. This study focused on DHFR variants from *Escherichia coli* which have exhibited resistance to trimethoprim. This study entailed a multi-prong approach that utilized nativespray MS and UVPD, size exclusion chromatography, and measurement of Michaelis-Menten kinetics and inhibitory kinetics by spectrophotometry. Results showed that the two mutants, P21L and W30R, resisted inhibition by TMP through two separate mechanisms. Whereas the W30R mutant relied on a reordering of the substrate-binding pocket for steric resistance, the rigidity of the M20 active loop was modulated for the P21L variant which allowed facile exchange between the substrate and TMP while retaining the substrate binding pocket structure.

In the final chapter, UVPD-MS was used to characterize the conformation of the Kirsten rat sarcoma (K-Ras) viral oncogene homolog and two single point mutations. K-Ras performs GTP hydrolysis in the cell upon interaction with its effectors in the cell. Mutations of K-Ras are very common in a wide range of cancers. These mutations impede the interaction with said effectors which in turn leaves K-Ras in the activated state with GTP bound. This leads to elevated GTP levels in the cell signaling for cell proliferation and downstream signaling. It was found that the structural dynamics or conformations changed significantly in their function with respect to the type of amino acid substitution at the Gly12 position in the K-Ras protein. These observations suggest

that specific G12X mutations may induce mutation-specific down-stream interactions driving cancer growth.

1.4 REFERENCES

(1) Kendrew, J. C.; Parrish, R. G. *Proc. R. Soc. Lond. Math. Phys. Eng. Sci.* **1957**, *238* (1214), 305–324.
(2) Shi, Y. *Cell* **2014**, *159* (5), 995–1014.
(3) Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. *FEBS J.* **2008**, *275* (1), 1–21.
(4) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci.* **2007**, *104* (23), 9615–9620.
(5) Wüthrich, K. *Nat. Struct. Mol. Biol.* **2001**, *8* (11), 923–925.
(6) Søndergaard, C. R.; Garrett, A. E.; Carstensen, T.; Pollastri, G.; Nielsen, J. E. *J. Med. Chem.* **2009**, *52* (18), 5673–5684.
(7) Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Mol. Biol.* **2002**, *319* (1), 209–227.
(8) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.
(9) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682.
(10) Rappsilber, J. *J. Struct. Biol.* **2011**, *173* (3), 530–540.
(11) Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. *Nat. Methods* **2015**, *12* (12), 1179–1184.
(12) Schmidt, C.; Robinson, C. V. *Nat. Protoc.* **2014**, *9* (9), 2224–2236.
(13) Suckau, D.; Mak, M.; Przybylski, M. *Proc. Natl. Acad. Sci.* **1992**, *89* (12), 5630–5634.
(14) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (15), 7391–7397.
(15) Zhou, Y.; Vachet, R. W. *Anal. Chem.* **2013**, *85* (20), 9664–9670.
(16) Zhang, H.; Liu, H.; Blankenship, R. E.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2015**, *27* (1), 178–181.
(17) Takahashi, K. *J. Biol. Chem.* **1968**, *243* (23), 6171–6179.
(18) Fliss, H.; Viswanatha, T. *Can. J. Biochem.* **1979**, *57* (11), 1267–1272.
(19) Yem, A. W.; Epps, D. E.; Mathews, W. R.; Guido, D. M.; Richard, K. A.; Staite, N. D.; Deibel, M. R. *J. Biol. Chem.* **1992**, *267* (5), 3122–3128.
(20) Zhou, Y.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (4), 708–717.
(21) Houston, L. L.; Walsh, K. A. *Biochemistry (Mosc.)* **1970**, *9* (1), 156–166.
(22) Horton, H. R.; Koshland, D. E. *J. Am. Chem. Soc.* **1965**, *87*, 1126–1132.
(23) Zhou, Y.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (5), 899–907.
(24) Madsen, J. A.; Yin, Y.; Qiao, J.; Gill, V.; Renganathan, K.; Fu, W.-Y.; Smith, S.; Anderson, J. *Anal. Chem.* **2016**, *88* (4), 2478–2488.
(25) Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C. *Anal. Chem.* **2012**, *84* (10), 4411–4418.
(26) Zhang, B.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2015**, *27* (3), 552–555.

(27) Gross, D. S.; Williams, E. R. *Int. J. Mass Spectrom. Ion Process.* **1996**, *157-158*, 305–318.
(28) Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L. *Anal. Chem.* **2011**, *83* (1), 311–318.
(29) Hambly, D.; Gross, M. *Int. J. Mass Spectrom.* **2007**, *259* (1–3), 124–129.
(30) Jones, L. M.; Sperry, J.; Carroll, J.; Gross, M. L. *Anal. Chem.* **2011**, *83* (20), 7657–7661.
(31) Hochleitner, E. O.; Borchers, C.; Parker, C.; Bienstock, R. J.; Tomer, K. B. *Protein Sci.* **2000**, *9* (3), 487–496.
(32) Kahraman, A.; Herzog, F.; Leitner, A.; Rosenberger, G.; Aebersold, R.; Malmström, L. *PLOS ONE* **2013**, *8* (9), e73411.
(33) Zybailov, B. L.; Glazko, G. V.; Jaiswal, M.; Raney, K. D. *J. Proteomics Bioinform.* **2013**, *6* (Suppl 2), 001.
(34) Chait, B. T.; Cadene, M.; Olinares, P. D.; Rout, M. P.; Shi, Y. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (6), 952–965.
(35) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. *Mol. Cell. Proteomics MCP* **2010**, *9* (8), 1634–1649.
(36) Merkley, E. D.; Rysavy, S.; Kahraman, A.; Hafen, R. P.; Daggett, V.; Adkins, J. N. *Protein Sci.* **2014**, *23* (6), 747–759.
(37) Pham, N. D.; Parker, R. B.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2013**, *17* (1), 90–101.
(38) Gomes, A. F.; Gozzo, F. C. *J. Mass Spectrom. JMS* **2010**, *45* (8), 892–899.
(39) Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. *Mol. Cell. Proteomics* **2011**, *10* (1), M110.002212.
(40) Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyö, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125* (9), 2416–2425.
(41) Engen, J. R. *Anal. Chem.* **2009**, *81* (19), 7870–7875.
(42) Baldwin, R. L. *Proteins Struct. Funct. Bioinforma.* **2011**, *79* (7), 2021–2026.
(43) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(44) McAllister, R. G.; Konermann, L. *Biochemistry (Mosc.)* **2015**, *54* (16), 2683–2692.
(45) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(46) Hayashi, N.; Kuyama, H.; Nakajima, C.; Kawahara, K.; Miyagi, M.; Nishimura, O.; Matsuo, H.; Nakazawa, T. *Biochemistry (Mosc.)* **2014**, *53* (11), 1818–1826.
(47) Cummins, D. J.; Espada, A.; Novick, S. J.; Molina-Martin, M.; Stites, R. E.; Espinosa, J. F.; Broughton, H.; Goswami, D.; Pascal, B. D.; Dodge, J. A.; Chalmers, M. J.; Griffin, P. R. *Anal. Chem.* **2016**, *88* (12), 6607–6614.
(48) Abzalimov, R. R.; Kaltashov, I. A. *Anal. Chem.* **2010**, *82* (3), 942–950.
(49) Hoerner, J. K.; Xiao, H.; Dobo, A.; Kaltashov, I. A. *J. Am. Chem. Soc.* **2004**, *126* (24), 7709–7717.
(50) Bobst, C. E.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (11), 5225–5231.

(51) Abzalimov, R. R.; Kaplan, D. A.; Easterling, M. L.; Kaltashov, I. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1514–1517.
(52) Wang, H.; Shu, Q.; Rempel, D. L.; Frieden, C.; Gross, M. L. *Biochemistry (Mosc.)* **2015**, *54* (42), 6475–6481.
(53) Zhang, Y.; Rempel, D. L.; Zhang, J.; Sharma, A. K.; Mirica, L. M.; Gross, M. L. *Proc. Natl. Acad. Sci.* **2013**, *110* (36), 14604–14609.
(54) Malito, E.; Faleri, A.; Surdo, P. L.; Veggi, D.; Maruggi, G.; Grassi, E.; Cartocci, E.; Bertoldi, I.; Genovese, A.; Santini, L.; Romagnoli, G.; Borgogni, E.; Brier, S.; Passo, C. L.; Domina, M.; Castellino, F.; Felici, F.; Veen, S. van der; Johnson, S.; Lea, S. M.; Tang, C. M.; Pizza, M.; Savino, S.; Norais, N.; Rappuoli, R.; Bottomley, M. J.; Masignani, V. *Proc. Natl. Acad. Sci.* **2013**, *110* (9), 3304–3309.
(55) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(56) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(57) Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683–693.
(58) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2012**, *85* (1), 2–9.
(59) Skinner, O. S.; McLafferty, F. W.; Breuker, K. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (6), 1011–1014.
(60) Schennach, M.; Breuker, K. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 164–168.
(61) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(62) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. *Nature* **2014**, *510* (7503), 172–175.
(63) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3* (2), 126–132.
(64) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(65) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16* (1), 1–23.
(66) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(67) Jurneczko, E.; Barran, P. E. *The Analyst* **2011**, *136* (1), 20–28.
(68) Niu, S.; Rabuck, J. N.; Ruotolo, B. T. *Curr. Opin. Chem. Biol.* **2013**, *17* (5), 809–817.
(69) Snijder, J.; Heck, A. J. R. *Annu. Rev. Anal. Chem.* **2014**, *7* (1), 43–64.
(70) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(71) Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. *Nat. Commun.* **2013**, *4*, 1985.
(72) Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 190–208.
(73) Zhou, M.; Wysocki, V. H. *Acc. Chem. Res.* **2014**, *47* (4), 1010–1018.

(74) Song, Y.; Nelp, M. T.; Bandarian, V.; Wysocki, V. H. *ACS Cent. Sci.* **2015**, *1* (9), 477–487.
(75) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(76) Hall, Z.; Politis, A.; Robinson, C. V. *Structure* **2012**, *20* (9), 1596–1609.
(77) Cong, X.; Liu, Y.; Liu, W.; Liang, X.; Russell, D. H.; Laganowsky, A. *J. Am. Chem. Soc.* **2016**, *138* (13), 4346–4349.
(78) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.
(79) Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V. *Nat. Methods* **2016**, *13* (4), 333–336.
(80) Bechara, C.; Nöll, A.; Morgner, N.; Degiacomi, M. T.; Tampé, R.; Robinson, C. V. *Nat. Chem.* **2015**, *7* (3), 255–262.
(81) McAlary, L.; Yerbury, J. J.; Aquilina, J. A. *Sci. Rep.* **2013**, *3*.
(82) Cubrilovic, D.; Haap, W.; Barylyuk, K.; Ruf, A.; Badertscher, M.; Gubler, M.; Tetaz, T.; Joseph, C.; Benz, J.; Zenobi, R. *ACS Chem. Biol.* **2014**, *9* (1), 218–226.
(83) Cubrilovic, D.; Barylyuk, K.; Hofmann, D.; Walczak, M. J.; Gräber, M.; Berg, T.; Wider, G.; Zenobi, R. *Chem. Sci.* **2014**, *5* (7), 2794–2803.
(84) Yin, S.; Loo, J. A. *Int. J. Mass Spectrom.* **2011**, *300* (2-3), 118–122.
(85) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(86) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2010**, *132* (25), 8602–8609.
(87) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(88) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(89) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(90) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72* (3), 563–573.
(91) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(92) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(93) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(94) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(95) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(96) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804–2812.

(97) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.
(98) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(99) Sawaya, M. R.; Kraut, J. *Biochemistry (Mosc.)* **1997**, *36* (3), 586–603.

Chapter 2: Experimental

2.1 Orbitrap Mass Spectrometer

A Thermo Orbitrap Velos Elite mass spectrometer was used for all analysis involving UVPD fragmentation. It has been modified for UVPD in the HCD cell as previously described in Shaw et al.[1] A schematic representation with labelled sections of the Orbitrap mass spectrometer are shown in Figure 1. The first section depicts the ESI source which generates ionized proteins which enter the heated capillary for further desolvation. After desolvation the ions are direct through ion optic multipoles to the dual pressure linear ion trap (section 3).

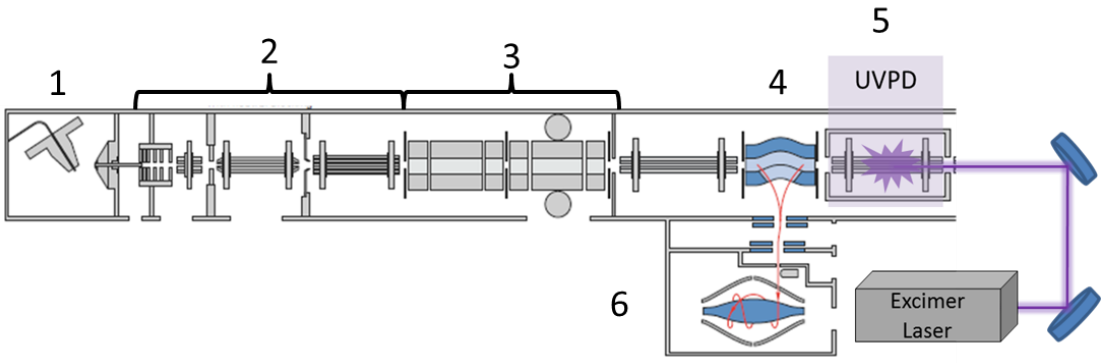

Figure 1. Schematic representation of Thermo Scientific Instruments Orbitrap Elite mass spectrometer illustrating the six main sections described in the text. The sections include: 1) ESI source, 2) ion optics, 3) dual pressure linear ion traps, 4) C-trap, 5) HCD cell where UVPD is implemented, 6) Orbitrap mass analyzer.

The dual pressure linear ion trap is responsible for low resolution mass detection as well as performing collisional activation and is also the section of the instrument that is capable of performing ETD and ion isolation. From here intact ions or those generated from fragmentation that require high resolution analysis are passed to section 4 of the

instrument deemed the C-trap. The C-trap is responsible for collecting the ions before injection into the orbitrap (section 6) for high resolution detection. Ions can also be fragmented with beam type collisional activation in the HCD cell (section 7). The HCD cell is also where UVPD occurs by irradiation of ions using a 193 nm Coherent excimer laser. Orbitrap detection generally allows intact proteins or fragments of proteins to be detected from *m/z* 200-4000 with resolution of 120k to 240k at 400 *m/z*. The orbitrap is a high accuracy detector yielding sub 10 ppm error tolerances for mass detection.

2.2 Protein Crosslinking Experimentation

2.2.1 Crosslinking Reactions

Crosslinking reactions are used to covalently link two reactive and accessible amino acids between two separate interacting proteins or within a single protein itself. Experiments such as these reveal spatial information about the protein interactions or the nearness of particular residues in one protein based on the crosslink distance constraints. Many different reactive crosslinking chemistries have been developed and are employed to study a wide array of protein interactions. Crosslinkers can also be synthesized with varying spanning lengths to discriminate between short and long range interactions.

Bovine ubiquitin and bovine insulin were prepared in a 1X PBS buffer prior to labeling with the homobifunctional crosslinker bis(sulfosuccinimidyl)suberate (BS3) (Figure 2). Zinc acetate was added as well for the insulin sample to induce hexamer formation. Proteins were reacted at a 1:10 ratio with BS3 (protein:BS3) (Figure 2) for 10 minutes to avoid excessive crosslinking. Samples were quenched by addition of 0.1%

formic acid. Excess BS3 was removed by passing the solution through a 3 kDa molecular weight cut-off filter.

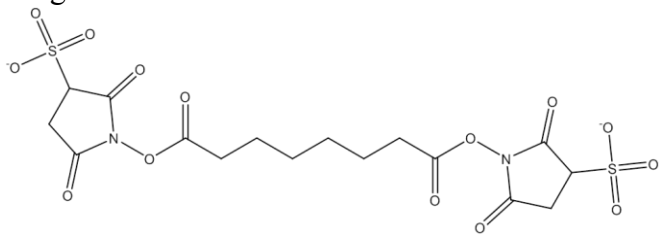

Figure 2. BS3 (bis(sulfosuccinimidyl)suberate) crosslinker

2.2.2 Reversed-Phase Liquid Chromatography

A Dionex nano-LC system was equipped with a 5 cm trapping column and a 30 cm analytical column each packed in house with polymeric reversed phase media (PLRP-S) at a 5μM particle size. The column had a 360 μM outer diameter with a 75 μM inner diameter. The trap was washed at 5 μL/min with loading solvent, composed of 98% water, 2% acetonitrile and 0.1% formic acid, after sample loading for a minimum of 6 minutes, then the trap column and analytical were aligned by a valve switch and the mixture of crosslinked proteins were separated at 400 nl/min. The solvent system used contained 0.1% formic acid in water (solvent A) or 0.1 % formic acid in acetonitrile (solvent B). Proteins were eluted over a gradient of 30 minutes from 20% to 40% solvent B. 100 ng of protein material was typically injected for each experiment. The nanoLC system was interfaced to an Orbitrap Elite mass spectrometer that was equipped with a 193 nm laser for UVPD. A 2kV potential was applied at the tip of the column to induce nanospray.

2.2.3 Targeted MS/MS Fragmentation

Targeted UVPD and HCD runs were performed on the singly intra-protein crosslink for ubiquitin and the singly inter-protein crosslinked insulin. For ubiquitin the

7+ species (*m/z* 1243.68) was targeted. For insulin the 8+ species (*m/z* of 1451.42) was targeted. All MS2 spectra were collected at 120,000 resolution at 400 *m/z*. A single 3 mJ pulse from the 193 nm laser was used for UVPD, and 31% NCE was applied for HCD fragmentation. 20-50 scans were averaged per MS/MS spectrum.

2.2.4 Data Analysis

MS2 scans were averaged across the most abundant and well-resolved peaks with Gaussian shape in comparison to the baseline. The spectra were deconvoluted by using Thermo Xtract with a S/N threshold of 3. The produced monoistopic peak list was analyzed with ProSight Lite, a database search algorithm and top-down data analysis algorithm originally developed in the Kelleher lab. For each decharged spectrum of ubiquitin, two separate searches were performed, one to identify the unmodified fragments, and the second to identify sites of crosslinking. The second was searched by adding +138. 069 Da (the mass addition of the crosslinker) to each termini to identify fragment ions that were covalently crosslinked. The sequence offset of the unmodified series of fragments and the onset of the modified fragments indicated the sites of crosslinking. The data was searched similarly for crosslinked insulin, but instead the summed mass of insulin plus the BS3 mass was added at all possible sites of crosslinking on the insulin A and B chains. There was one possible site on the A chain at the N-terminus and two sites on the B chain at the N-terminus and K29 (or KB). As with ubiquitin the resulting series of unmodified and modified fragments indicated the sites of crosslinking within insulin dimers. Note that the crosslinked dimers were produced from a hexameric form of insulin in solution.

2.3 Native Mass Spectrometry of Proteins

Native mass spectrometry of proteins uses non-denaturing conditions in contrast to standard denaturing ESI used for protein analysis. These conditions are most dependent upon using volatile salt buffers such as ammonium acetate or ammonium bicarbonate at biologically relevant pH. This is in contrast to typical denaturing protein ESI which the solvent is composed of greater than 50% organic solvent (methanol or acetonitrile) with the addition of low percentages of acetic or formic acid. When applicable a static tip emitter spray is used instead of an ESI source. This allows the application of very low voltages (0.8 – 2kV) for protein ionization in comparison to standard ESI (3-4kV). Using substantially lower voltages in turn lowers the internal energy of ions based on the heat generated in the resulting droplet produced either by ESI or through a static tip setup. It is also typical to lower the heated capillary temperature from 300°C (for denaturing ESI) to a range between 50°C to 200°C for native experiments. The overall goal of spraying a native protein is to keep the internal energy of the gaseous protein as low as possible while it travels into and through the mass spectrometer by adjusting said parameters above. This low energy approach is done at the expense of sensitivity.

2.3.1 Mass Spectrometry

All native mass spectrometry experiments were undertaken using an Orbitrap Elite mass spectrometer which was interfaced to a 193 nm excimer laser for UVPD. All experiments using native mass spectrometry were performed by static spray using an offline nanoESI kit to modify the existing nanoESI source. This nanoESI kit consisted of a ring electrode to hold and apply voltage to the coated tips along with the ability to apply backpressure to the static tip during ionization. 1.2 mm outer diameter borosilicate capillaries were pulled to fashion 20-50 μM inner diameter tips, then coated to 20 nm

thickness uniformly on the capillary surface with a gold:palladium (20:80) mixture. These coated borosilicate capillaries were used for all nativespray experiments. In this fashion the ring electrode from the offline nanoESI source came in contact with the metal on the tip and induced an electroosmotic current which caused the sample to flow and resulted in production of a spray. Spray voltages between 0.8 to 2.0 kV were used for spraying the proteins. All proteins were sprayed out of varying concentrations of ammonium acetate solvent and pH values. K-Ras proteins were specifically sprayed out of 50 mM ammonium acetate at pH 7.8 to facilitate full co-factor binding (Chapter 6). DHFR proteins were sprayed out of 50 mM ammonium acetate at pH 6.5 in Chapter 4 and 150 mM ammonium acetate at pH 6.5 with 2% DMSO to facilitate anti-folate solubility in solution in Chapter 5. The temperature of the heated capillary (a component of the ESI interface that facilitated desolvation of ions) remained between 100 – 200°C for all native experiments. MS2 of each protein was performed using a single 3 mJ pulse of the 193 nm laser. MS/MS spectra were acquired at a resolution of 120k at 400 *m/z*, and up to 250 scans were averaged for DHFR experiments, and a resolution of 240k at 400 *m/z* and up to 750-1000 scans were averaged for K-Ras experiments. The laser power was adjusted for DHFR analysis in Chapter 5 to 2.5 mJ. For the acquisition of UVPD mass spectra, the Orbitrap was scanned from *m/z* 220 to 4000 with a standard 10 mTorr pressure in the HCD cell.

2.3.2 Size Exclusion Chromatography – Native MS

SEC experiments were performed on a Dionex LC system interfaced with a Thermo Velos Pro dual cell ion trap mass spectrometer. The LC was interfaced with the mass spectrometer using a HESI source. The HESI source differs from the nanoESI source in that typical spray voltages range from 3-4kV and nitrogen sweep gas and

nitrogen auxiliary gas can be used to improve desolvation. A HESI source is used for capillary to analytical flow experiments. The Dionex LC was set up in a direct injection fashion bypassing any preconcentration steps. In short the loading pump was plumbed through the autosampler and then directly to the entrance of the column and then attached directly to the HESI source. A 4 kV voltage was applied at the source to facilitate ESI. A 2.1 mm x 30 cm Zenix-C column with an 80 Å pore size (80,000 MW exclusion limit, protein MW range 100-80,000 Da) and 3 μm particle size from Sepax Technologies was employed for SEC separation. The stationary phase was hydrophilic film bonded silica. The mobile phase for all SEC experiments was 150 mM ammonium acetate at pH 6.5 which flowed at 80 μL/min. Each protein solution was composed of 12 μM DHFR with excess ligands in 150 mM ammonium acetate with 2% DMSO. 5 μg of protein was injected for each experiment (which corresponded to an injection volume of 20 uL of the aforementioned solution). The mass spectrometer was set to scan from *m/z* 1500-4000 and a small 10 *m/z* window centered around the expected ligand(s) *m/z* in an alternating fashion to allow monitoring the protein complexes and free ligands.

2.3.3 Data Analysis

All collected UVPD mass spectra were deconvoluted to monoisotopic masses by using Xtract with a S/N of 2 to 3 prior to data analysis. Chapters 4 and 6 used Prosight PC 3.0, from the Kelleher lab, that was modified to accompany searching UVPD type ions (a, a•, b, c, x, x•, y, y-1, and z). Chapter 5 involved the use of a fragment identification algorithm developed in-house by Jake Rosenberg. The in-house algorithm produced a list of identified ions with their respective intensities. It also allowed facile summation of ion intensities over the entire backbone as well as allowed this information to be split into summed ion intensities arising from the N versus C terminus.

Additionally, the algorithm offered the ability to search for numerous mass shifts induced by ligand binding. These two described utilities of automatic ion summation and mass shifts due to ligand binding are not available with the current commercially available top down software (Prosight PC 3.0). In general two types of searches were performed for each protein-ligand complex. The first revealed the ion abundances of apo (ligand free) fragment ions. The second revealed the holo (ligand-containing) fragment ion abundances by searching all fragment ions again while including the expected mass of the ligated small molecules added on both the N- and C-termini of the protein. To normalize, the abundances of the identified fragment ions were divided by the total ion current of the recorded mass spectra. The N- and C-terminal ions were then summed in a complementary fashion across the backbone from the N terminus to C terminus. For example, N-terminal ions with fragment position n (a_n, b_n, c_n) are summed with C-terminal ions with fragment position n or (x_{R-n+1}, y_{R-n+1}, z_{R-n+1}) where R is total number of residues in the protein. (Figure 3) These two series of ions (apo and holo) were then summed together to generate the total fragment ion abundance across the protein backbone.

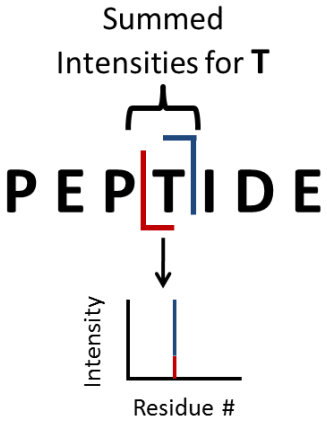

Figure 3. Pictorial description of summing N and C terminal ion intensities per amino acid residue.

2.4 REFERENCES

(1) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

Chapter 3: Structural Characterization of Intra- and Intermolecular Protein Crosslinking by Top Down Ultraviolet Photodissociation Mass Spectrometry

3.1 Overview

Interactions within and between proteins impact structure and function, and mapping these interactions is a key hallmark of structural proteomics. One popular method for mapping protein interactions utilizes homobifunctional crosslinkers with defined distance constraints, followed by bottom-up mass spectrometry analysis. In this study, characterization of protein crosslinks was accomplished by top-down mass spectrometry (MS) using both ultraviolet photodissociation (UVPD) and higher energy collisional dissociation (HCD) in conjunction with reversed phase nanoscale liquid chromatography (nanoLC). Four intramolecular crosslinks of ubiquitin were identified, all in agreement with the known tertiary structure. Three intermolecular crosslinks of insulin were unambiguously assigned, consistent with the hexameric complex adopted by insulin in solution. This integrated top-down nanoLC/UVPD/HCD-MS approach affords a powerful strategy for deciphering details about tertiary structure and intermolecular protein interactions.

3.2 Introduction

Chemical crosslinking of proteins and protein complexes followed by proteolysis and analysis of the crosslinked peptides by conventional bottom-up mass spectrometry workflows has proven to be an exceptionally powerful strategy for providing key insight into protein conformations, protein interfaces, and protein interactomics.[1] However, by their very nature bottom-up methods create very complicated mixtures, ones in which low abundance cross-linked peptides might be missed, and each crosslinked peptide

provides at most one contact point. A compelling alternative is a top-down approach in which intact proteins (or protein complexes) are analyzed to provide a more holistic perspective for deciphering protein interactions.[2,3] Recent developments of new ion activation techniques[4,5] and ongoing improvements of high resolution instrumentation[6] have advanced the field of top-down mass spectrometry to allow greater sequence coverage of individual proteins[7] and grander scale mapping of the whole proteome.[8] In terms of top-down strategies for crosslinked proteins, only two studies have been reported,[9,10] both focusing on characterization of ubiquitin crosslinked in solution or in the gas phase, respectively, by homobifunctional amine-specific N-hydroxysuccinimide ester-type linkers. Each study used collisional activated dissociation for characterization of the crosslinked ubiquitin, and each found two crosslinks, K6 to K11 and K48 to K63[9] or K27 to K29 and K48 to K63,[10] respectively. The studies in which crosslinking was combined with top-down MS/MS analysis are over a decade old[9,10] and have not been surpassed with newer findings, thus attesting to the challenges of top-down approaches for examination of crosslinked proteins. The primary limitations are apparent: sub-par sequence coverage from conventional collisional activation which restricts confident localization of crosslinks, as well as the difficulty of deriving site-specific information from heterogeneous populations of crosslinked proteins. We have addressed these hurdles by employing both ultraviolet photodissociation (UVPD) and higher energy collisional dissociation (HCD) to provide an unprecedented level of sequence coverage in conjunction with nanoscale liquid chromatography (nanoLC) to separate the crosslinked proteins, ultimately allowing unambiguous identification of the crosslinks. This strategy is demonstrated for intramolecular crosslinked ubiquitin and intermolecular crosslinked hexameric insulin.

3.3 EXPERIMENTAL

3.3.1 Crosslinking of Ubiquitin and Insulin

Bovine ubiquitin (50 μM) or bovine insulin (15 μM) were prepared in 1X PBS buffer, the latter with additional zinc acetate added to five times the insulin concentration to promote formation of the active insulin hexamer. Each solution was reacted with bis(sulfosuccinimidyl)suberate) (BS3) at a 1:10 (protein:BS3) molar ratio for 10 minutes, then subsequently quenched via addition of approximately 5 μL of 0.5% formic acid. The protein:BS3 ratio and reaction time were kept low to minimize formation of multiple crosslinks per protein, an outcome that confounds top-down MS/MS spectra interpretation. Excess BS3 was removed by passing each solution through a 3 kDa molecular weight cut-off filter.

3.3.2 Mass Spectrometry and Liquid Chromatography

100 ng of protein material was injected on a nanoLC system equipped with a 5 cm loading polymeric reverse phase media (PLRP-S) column at 5 uL/min prior to a 30 cm PLRP-S column at 400 nl/min with 0.1% formic acid in water (A) and acetonitrile (B) for mobile phases. The nanoLC system was interfaced to a Thermo Scientific Orbitrap Elite mass spectrometer modified for UVPD in the HCD cell, as described previously.(7) The electrospray voltage was held at 1.8 kV, and the mass spectrometer was set to target intramolecular crosslinked ubiquitin (Mr 8697.68 Da, target m/z 1244.39 (7+)) or intermolecular crosslinked insulin (Mr 11597.29 Da, target m/z 1451.52 (8+)). The collision energy for HCD was set to NCE 31%. A single 3 mJ laser pulse was used for UVPD. All MS2 spectra were collected at 120 K resolution at m/z 400 averaging 5 μscans.

3.3.3 Data Analysis

MS/MS spectra were deconvolved using Xtract with a S/N ratio of 3. The resulting deconvoluted mass spectra were searched using ProSight Lite both with and without the crosslinking mass shift (+138.069 Da) fixed on either termini to determine sites of crosslinker attachment with a 6 ppm mass error tolerance. Inter-protein crosslinks for insulin were searched by fixing the mass of an insulin monomer plus the mass of the crosslinker (Mr 5867.69 Da) to the N-terminus of the A and B chains of insulin as well as K29 on the B chain.

3.4 Results and Discussion

3.4.1 Ubiquitin Crosslinks

The chromatographic profiles obtained for the BS3-crosslinked ubiquitin mixture are shown in **Figure 1a** with the unmodified, dead-end (i.e. hanging linker with one end hydrolysed, +156.083 Da), and crosslinked (+138.068 Da) species color-coded based on the extracted ion chromatograms for the 7+ species of each. Each type of product was successfully separated, and isomeric crosslinked species appeared as distinctive chromatographic peaks. Examination of the charge state distributions of each product type (shown in the inset of **Figure 1a**) revealed a decrease in average charge state going from unmodified ubiquitin to the dead-end product to the crosslinked species. This trend paralleled the number of basic sites; BS3 reacts with primary amines, converting them to less basic sites. Each of the crosslinked products was characterized by UVPD and HCD. The resulting MS/MS spectra for the 7+ charge states are shown in **Figure 2**, and the corresponding fragmentation maps are shown in **Figure 3**. Crosslinks were determined by locating the transition from unmodified fragment ions to fragments containing a mass shift corresponding to the crosslinker mass. The primary amines (lysine side-chain or N-

termini) closest to each transition were considered to be the sites that were crosslinked unless otherwise stated.

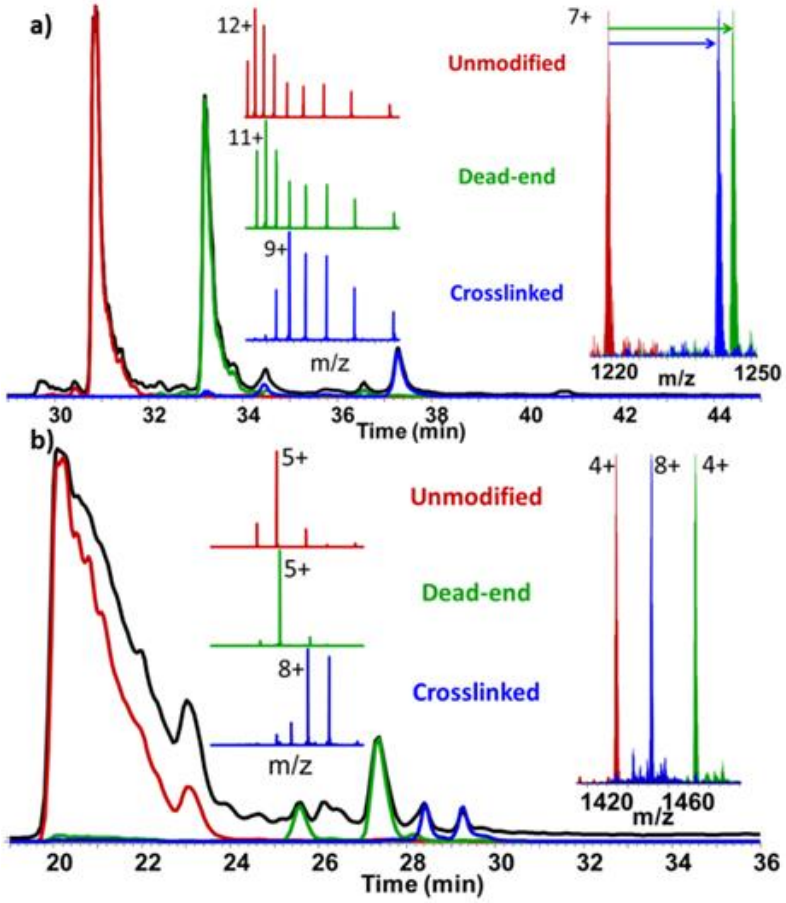

Figure 1. Chromatographic profile for a) crosslinked ubiquitin and b) hexameric insulin showing the total ion chromatographic trace (black) and extracted ion chromatograms for unmodified (red), dead-end (green) and crosslinked (blue) proteins. ESI-mass spectra from each species (middle) and expanded regions (far right, normalized and overlaid on a single m/z axis) of the mass spectra are shown next to the chromatographic profiles.

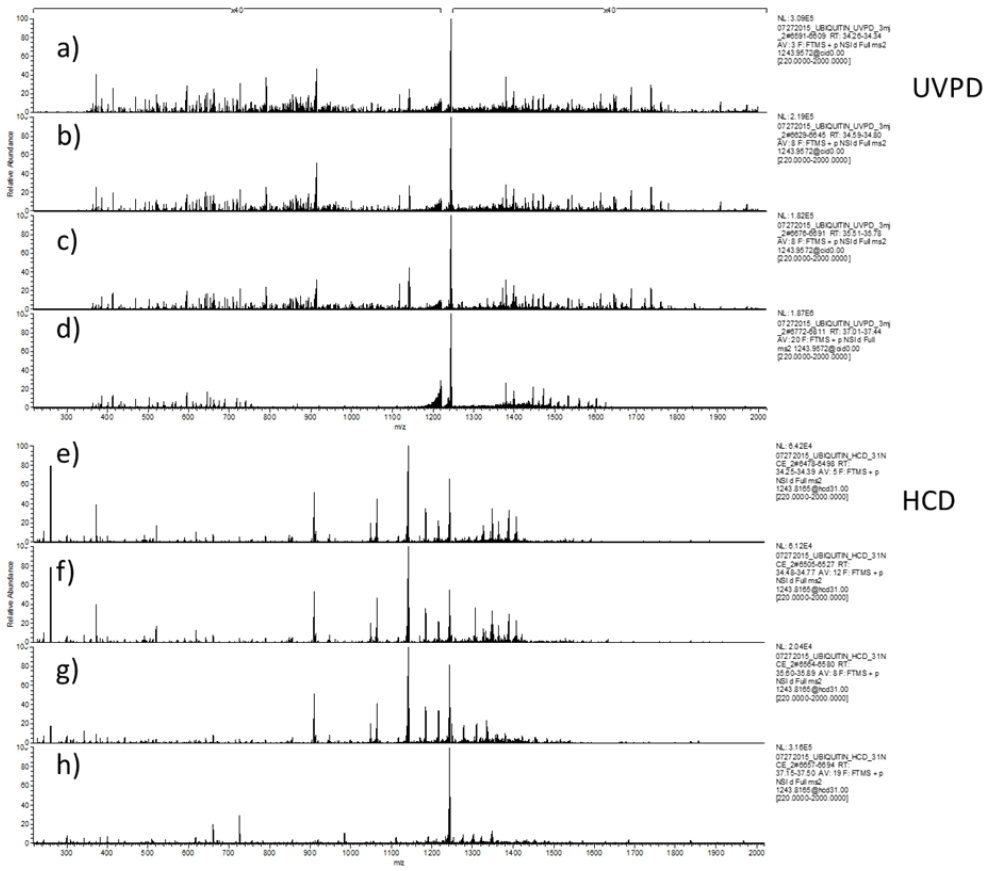

Figure 2. UVPD (a-d) and HCD (e-h) spectra of each crosslink of ubiquitin based on activation of the 7+ charge state. UVPD was performed using 1 pulse of 3 mJ, and 31% NCE was used for HCD.

Figure 3. Fragmentation maps (as designated by slash marks along the backbone) for UVPD (all on left) and HCD (all on right) for each crosslinked species of ubiquitin. For each of the four crosslinked species, a pair of fragmentation maps is shown, with the upper one representing unmodified fragment ions that don't contain the crosslink mass shift and the lower one reflecting fragment ions containing the crosslink mass shift (+138.0686 Da). For the lower maps in each pairing, the N-term and C-term residues are shaded orange to indicate that the crosslink mass shift was added to all N-term and C-term fragment ions during the searches. Yellow-shaded boxes outlined in blue indicate the sites of the crosslinks. Black boxes around some of the residues are non-informative and an artifact of the ProSight Lite display.

Analysis of the UVPD fragmentation patterns revealed three crosslinked species in which the following pairs of sites were linked by BS3: K11 to K33 (tr 34.3 min), N-terminus to K29 (tr 35.6 min) , and N-terminus to K63 (tr 37.2 min). Analysis of the complementary HCD spectra indicated that the first eluting crosslinked species was comprised of two products (**Figure 4a**, red and green extracted ion chromatograms). The second co-eluting crosslinked product (tr 34.6 min) was identified by the appearance of

an unmodified y58 fragment ion (m/z 1306.506, 5+ charge state), suggesting a crosslink between K6 and K11.

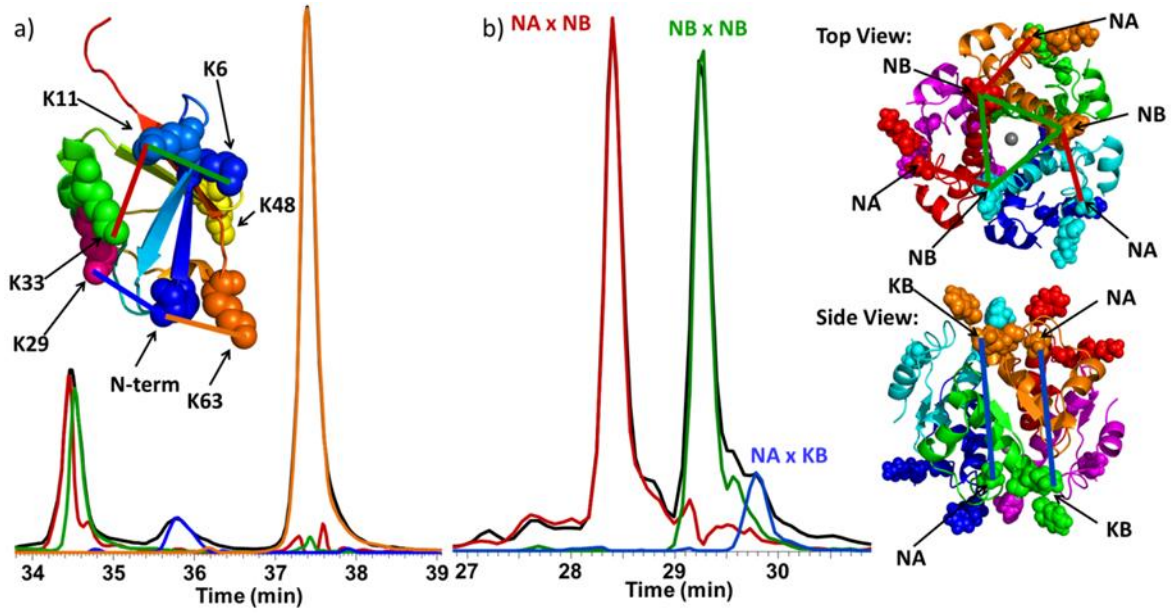

Figure 4. Extracted ion chromatograms (XIC) of the identified crosslinks in: a) ubiquitin (expanded chromatographic profile (t = 34 to 38 min) from Figure 1a) and b) hexameric insulin (expanded chromatographic profile (t = 27 to 31 min) from Figure 1b). Four intramolecular crosslinks were identified for ubiquitin and are shown as color coded bars in the NMR structure (PDB: 1D3Z) that match the colors of the XICs. Ions used for XIC of ubiquitin were m/z 1407.6 (y_{74} + xlink), 1307.3 (y_{58}), 1735.5 (b_{30} + xlink), and 1683.9 (internal fragment) for the red, green, blue and orange traces, respectively. Three intermolecular crosslinks were identified for hexameric insulin and are shown as color coded bars in the NMR structure (PDB: 1AI0) with a view of the top and side of the hexamer. The ions used for XIC were m/z 1475.18 (A-chain, a_1 + xlink + insulin), 526.28 (A-chain, b_5), and 1237.9 (B-chain, y_3 + xlink + insulin) for the red, green and blue traces, respectively.

Closer inspection of the extracted fragment ion abundances across the elution profile of the first crosslinked peak (tr 34.2-34.9 min) showed a demarcation among some of the fragment ion abundances, with a unique series of fragment ions occurring for the later-eluting species (green trace) suggesting a second isomeric crosslinked form (see

Figure 5). UVPD and HCD proved to be complementary because the N-terminus to K29 crosslink could not be confidently identified from the HCD fragmentation map, and the second crosslinked product (K6 to K11) was obscured in the richer UVPD fragmentation map.

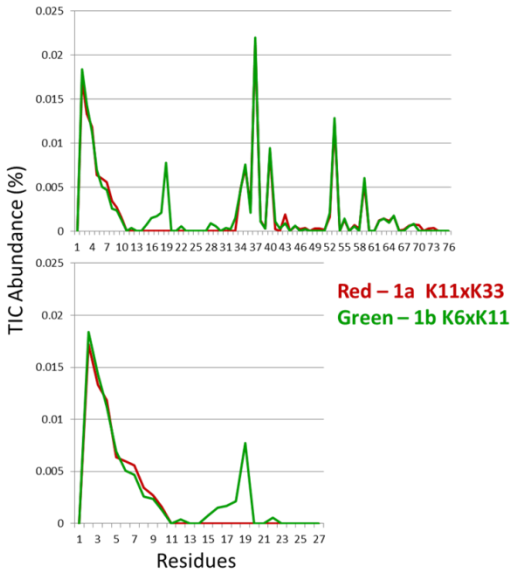

Figure 5. Overlay of backbone cleavage maps generated upon HCD of crosslinked ubiquitin (t_r 34.3 min) in which all N-terminal and C-terminal product ions attributed to cleavage at each backbone position are summed. The red traces are from peak 1a while the green traces are from peak 1b. These correspond to the traces in Figure 2 for the first co-eluting crosslinked species.

All crosslinks identified based on the UVPD and HCD spectra fell within the distance constraint of BS3 (maximum 24.2 Å spatial distance between Cα of each linked residue). Additionally, the extracted ion chromatographic profiles in **Figure 4** facilitated comparison of the relative abundances of the crosslinked products. The N-term to K63 crosslink was the most abundant and thus was considered the dominant crosslink formed. Interestingly, this particular crosslink cannot be successfully identified without chromatographic separation due to the convolution of fragment ions arising from

isomeric crosslinked species which prevents localization of the crosslink site. The expected K48 to K63 crosslinked product was not identified in the present study; apparently its low abundance prohibited detection and perhaps was due to inefficient reaction at K48 or insufficient chromatographic resolution to separate this crosslink and alleviate ionization suppression from other more abundant co-eluting products. Comparison of the number of diagnostic fragment ions generated by UVPD and HCD (**Figure 6**) indicated that UVPD generated significantly more fragment ions per crosslinked species, including both unmodified fragments (i.e. fragments that did not contain the crosslink) as well as fragments that incorporated the crosslink.

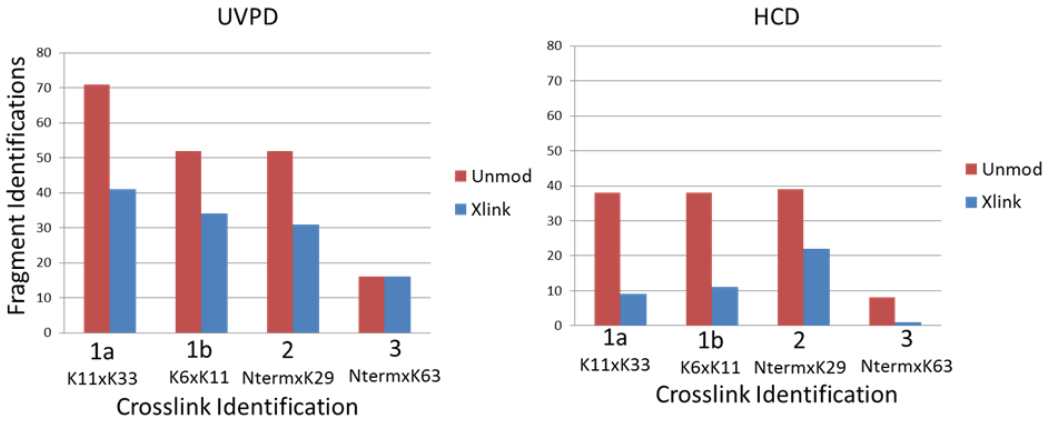

Figure 6. Fragment identifications per crosslink identification for UVPD and HCD based on searching for both unmodified and crosslinked fragments.

Interestingly, HCD produced a smaller portion of fragment ions that contained the crosslink and instead favored formation of unmodified fragment ions (N-terminal and C-terminal fragments from either side of the crosslinked segment). The identification of crosslinked sites provides spatial constraint information but does not allow construction of complete three dimensional structures.

3.4.2 Insulin Crosslinks

A similar story emerges for the crosslinks identified upon analysis of hexameric insulin (e.g. a trimer of dimers). Briefly, first the 5+ charge state of intact monomeric insulin (Mr 5729.62 Da) was characterized by UVPD and HCD to evaluate sequence coverages for a heavily disulfide-bound protein (**Figure 7**).

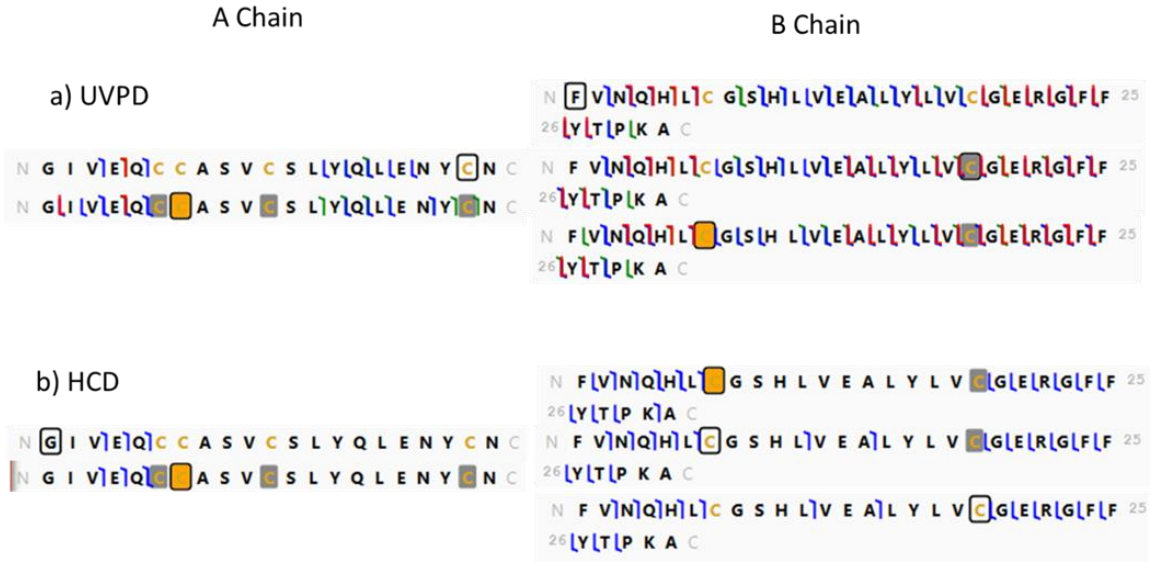

Figure 7. Fragmentation maps for both chains of intact monomeric insulin from (a) UVPD and (b) HCD. Each map represents a variation of the A chain or B chain depending on which disulfide bonds are cleaved. Grey-shaded cysteines indicate hydrogen losses to compensate for disulfide bonds formed in solution. Orange-shaded cysteines indicate addition of either the A chain (+2333.94 Da) or B chain (+3395.65 Da) of insulin with respect to the B or A chain, respectively. Black boxes around some of the residues are non-informative and an artifact of the Prosight Lite display.

UVPD gives extensive sequence coverage well beyond the disulfide-bound sections in comparison to the rather scanty interior fragmentation seen with HCD. Insulin forms hexameric complexes when in solution with zinc ions (**Figure 8**). After incubation of insulin with BS3, the resulting chromatographic profile highlighting unmodified monomeric insulin (4+), dead-end modified monomeric insulin (4+), and crosslinked

insulin (8+) is shown in Figure 1b. Each molecule of insulin has three primary amines: the N-terminus of the A chain (NA), the N-terminus of the B chain (NB), and Lys29 of the B chain (KB).

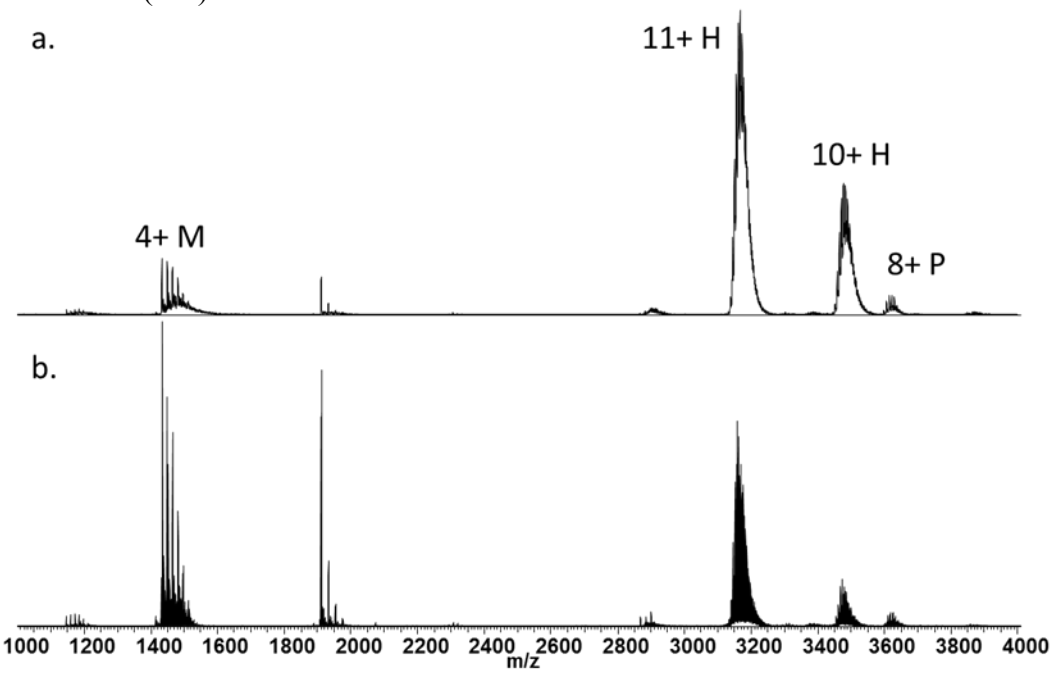

Figure 8. Nativespray mass spectra of hexameric insulin acquired (a) in the ion trap and (b) in the Orbitrap (240 k resolution at m/z 400. A 1.5 kV was applied to the tip. The solution contained 10 0mM ammonium acetate, 20 μM Zn acetate, and 10 μM bovine insulin. The monomeric (M) and hexameric (H) species the dominant species; there is some pentamer (P) present as well. Dimers are not observed.

This set of reactive sites affords up to six theoretical crosslink sites for any pair of insulin molecules, all within the expected 30 Å distance constraint of BS3.(11) Intermolecular crosslinks were elucidated by careful inspection of fragment ions bracketing the three possible sites of crosslinking (NA, NB, KB) based on the extensive UVPD fragmentation. As with ubiquitin, targeted UVPD and HCD were performed on the intermolecular crosslinked species of insulin (Mr 11597.29 Da, target m/z 1451.52 (8+)),

revealing two major species and one minor species (**Figure 4b**) via extracted fragment ion profiles. The first eluting product was crosslinked between NA and NB of two insulin molecules based on the fragmentation pattern from UVPD (**Figure 9, 10, and 11**).

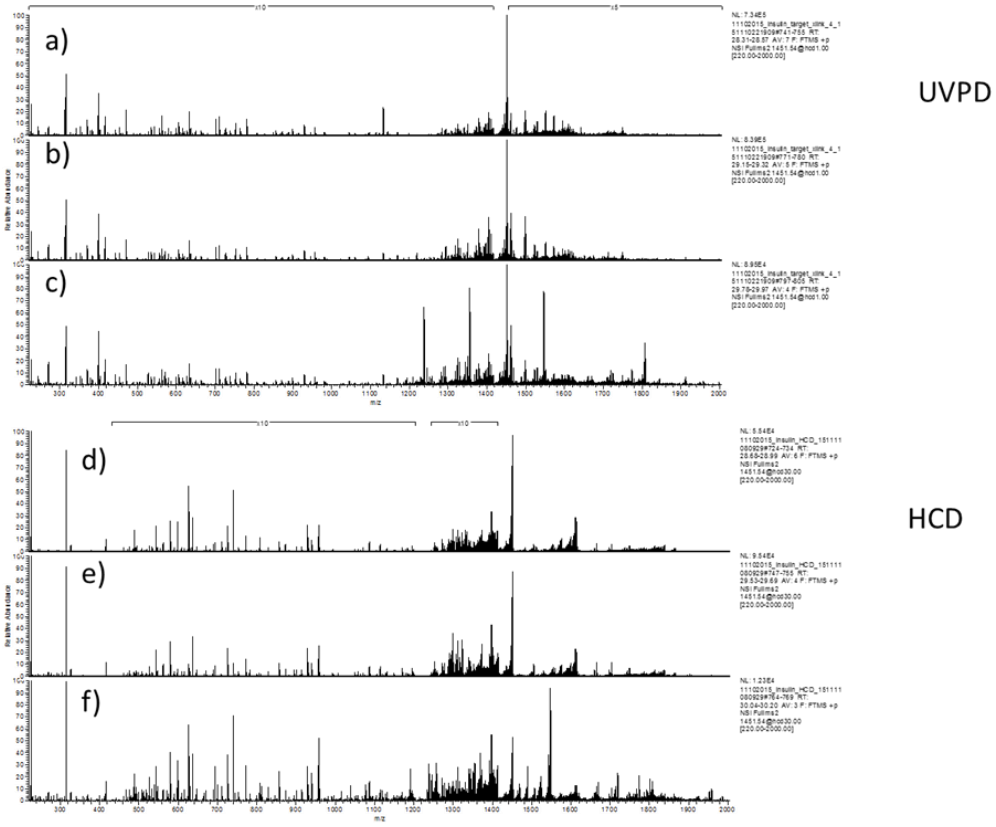

Figure 9. UVPD (a-c) and HCD (d-f) spectra of each of the three intermolecular crosslinked insulin forms based on activation of the 8+ charge state. UVPD was performed using 1 pulse of 3 mJ, and 31% NCE was used for HCD.

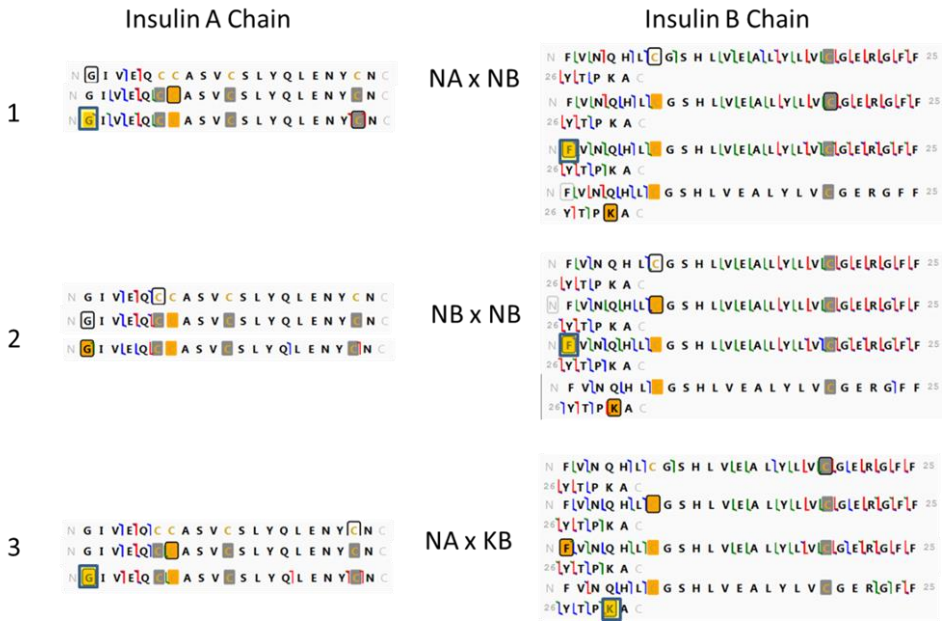

Figure 10. Fragmentation maps for both chains of insulin obtained by UVPD. Grey-shaded cysteines indicate hydrogen loss modifications to accommodate for disulfide bonds. Orange-shaded cysteines indicate addition of either the A (+2333.94 Da) or B (+3395.65 Da) chain of insulin with respect to the B or A chain, respectively. Crosslink modifications were searched at the N-terminus of the B chain (Phe residue) or K29 of the B chain as well as at the N-terminus of the A chain (Gly residue) with the addition of +5867.66 Da (corresponding to insulin + crosslinker mass). Yellow-shaded boxes outlined in blue indicate the sites of the crosslinks. There is only one yellow box in NBxNB. Black boxes are uninformative and are part of Prosight Lite display.

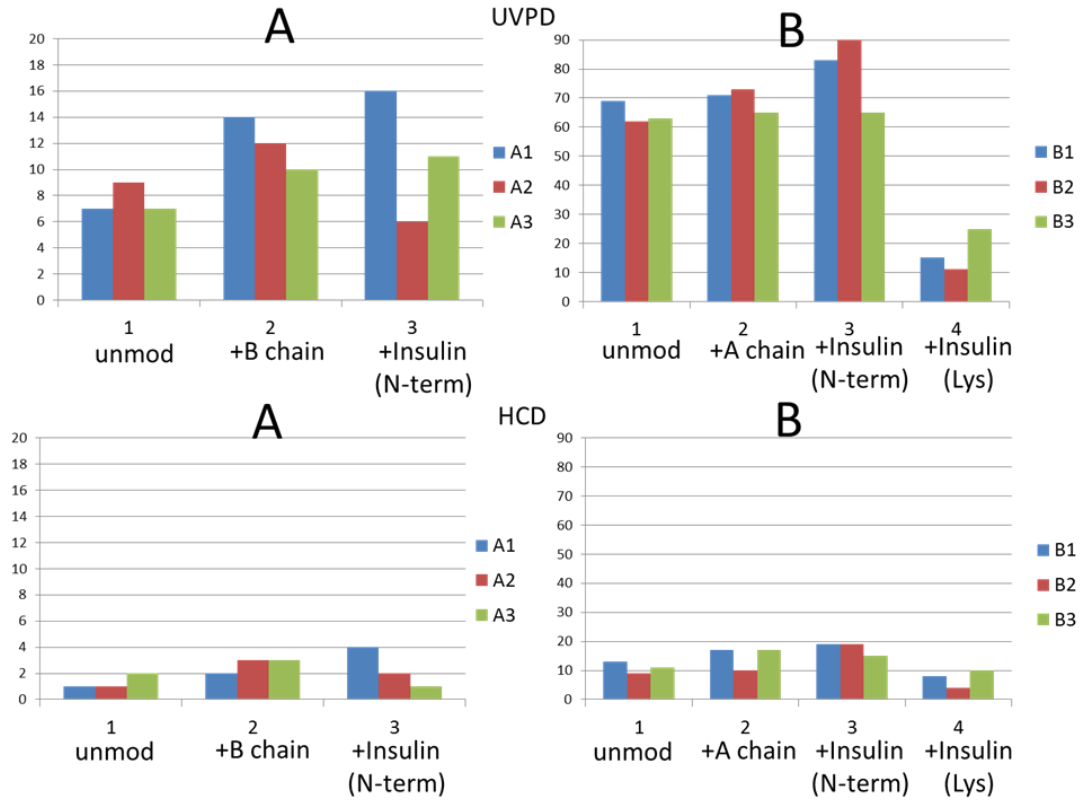

Figure 11. Number of fragment ions identified per search for each intermolecular crosslink species from hexameric insulin based on UVPD and HCD with respect to each chain of the protein. A1-3 or B1-3 indicate the chain as well as the specific LC peak from which the fragment identifications correspond. Peaks 1-3 correspond to NAxNB, NBxNB, and NAxKB, respectively. +B chain or +A chain indicate the addition of the mass of the A or B chain, respectively, to the indicated cysteine in Supplemental Figures 8 or 10. The addition of the crosslinked insulin mass is signified by +insulin at each point of possible attachment which include the N-terminus of both the A and B chains as well as the B chain lysine.

This crosslink assignment was confirmed by identification of fragment ions that matched N- or C-terminal sequence ions containing a mass shift corresponding to addition of one molecule of insulin plus the crosslinker (+5876.662 Da mass shift). For the second eluting crosslink species (NBxNB), there were few fragments ions that contained the modification on the N-terminus of the A chain, and there were a number that contained

the modification of the N-terminus of the B chain. The third eluting crosslink species was a minor product crosslinked between NA and KB based on the identification of many new crosslink-modified C-terminal ions from the B-chain and crosslink-modified N-terminal A-chain ions not observed previously. Additional crosslinked fragment ions related to the N-terminus of the B chain were due to chromatographic overlap from the second crosslinked product and thus were not diagnostic for the NAxKB crosslink. For crosslinked insulin, HCD did not provide any additional information nor did it contradict the assignments based on the UVPD spectra (**Figure 12**).

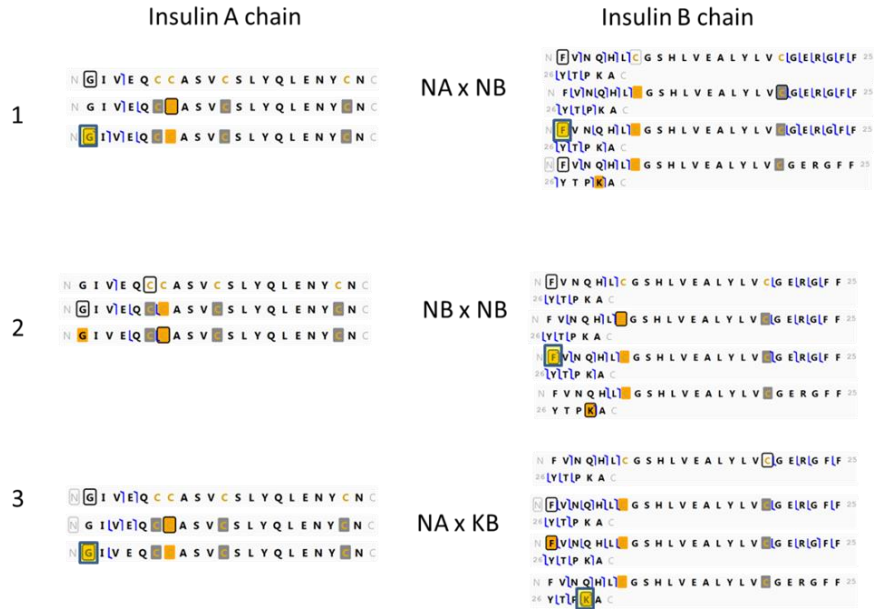

Figure 12. Fragmentation maps for both chains of insulin obtained by HCD. Grey-shaded cysteines indicate hydrogen loss modifications to accommodate for disulfide bonds. Orange-shaded cysteines indicate addition of either the A (+2333.94 Da) or B (+3395.65 Da) chain of insulin with respect to the B or A chain, respectively. Crosslink modifications were searched at the N-terminus of the B chain (Phe residue) or K29 of the B chain as well as at the N-terminus of the A chain (Gly residue) with the addition of +5867.66 Da (corresponding to insulin + crosslinker mass). Yellow-shaded boxes outlined in blue indicate the sites of the crosslinks. There is only one yellow box in NBxNB. Black boxes are uninformative and are part of Prosight Lite display.

3.5 CONCLUSION

We have demonstrated the ability to confidently characterize both intramolecular and intermolecular protein crosslinks by using a powerful top-down strategy that combined nanoLC with two complementary activation methods, HCD and UVPD. Identification of several minor and predominant crosslinked products was accomplished based on examination of chromatographic peak profiles and extensive sequence maps. The construction of extracted ion chromatographic profiles generated from structure-specific fragment ions allowed co-eluting crosslinked products to be resolved and characterized individually, a significant advantage given the complexity of fragmentation patterns of intact proteins. At this point conventional bottom-up methods provide greater sensitivity and better detection limits than top-down approaches, although in the context of identifying crosslinks the bottom-up LCMS/MS methods continue to be confounded by the difficulty of pinpointing low abundance crosslinked peptides in a complicated milieu of unmodified peptides. As algorithms for interpreting the fragmentation patterns of intact crosslinked proteins improve, the potential to examine species containing multiple simultaneous crosslinks will support construction of more comprehensive maps of integrated interactions and more detailed conformational insight. In addition, the possibility of examining interactions of proteins containing post-translational modifications, although challenging, represents a compelling opportunity arguably best addressed by top-down methods that can track combinatorial patterns of modifications.

To expand this top-down UVPD methodology to crosslinking applications involving larger protein systems, significant improvements in chromatographic resolution are needed to separate isomers, possibly by employing a two dimensional strategy such as ion exchange chromatography followed by reversed-phase LC. Additionally, HCD of

larger proteins typically yields low sequence information. More focus must be placed on implementing UVPD to accomplish larger crosslinked samples. This approach offers great promise for addressing increasingly elaborate questions about protein-protein interactions.

3.6 REFERENCES

(1) A. Sinz, *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
(2) A. D. Catherman, O. S. Skinner, N. L. Kelleher, *Biochem. Biophys. Res. Commun.* **2014**, *445*, 683–693.
(3) H. Zhou, Z. Ning, A. E. Starr, M. Abu-Farha, D. Figeys, *Anal. Chem.* **2012**, *84*, 720–734.
(4) J. S. Brodbelt, *Chem. Soc. Rev.* **2014**, *43*, 2757–2783.
(5) A. W. Jones, H. J. Cooper, *Analyst* **2011**, *136*, 3419–3429.
(6) M. E. Belov, E. Damoc, E. Denisov, P. D. Compton, S. Horning, A. A. Makarov, N. L. Kelleher, *Anal. Chem.* **2013**, *85*, 11163–11173.
(7) J. B. Shaw, W. Li, D. D. Holden, Y. Zhang, J. Griep-Raming, R. T. Fellers, B. P. Early, P. M. Thomas, N. L. Kelleher, J. S. Brodbelt, *J. Am. Chem. Soc.* **2013**, *135*, 12646–12651.
(8) A. D. Catherman, K. R. Durbin, D. R. Ahlf, B. P. Early, R. T. Fellers, J. C. Tran, P. M. Thomas, N. L. Kelleher, *Mol. Cell. Proteomics* **2013**, *12*, 3465–3473.
(9) G. H. Kruppa, J. Schoeniger, M. M. Young, *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155–162.
(10) I. K. Webb, M. Mentinova, W. M. McGee, S. A. McLuckey, *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 733–743.
(11) E. D. Merkley, S. Rysavy, A. Kahraman, R. P. Hafen, V. Daggett, J. N. Adkins, *Protein Sci.* **2014**, *23*, 747–759.

Chapter 4: Structural Characterization of Dihydrofolate Reductase Complexes by Top-down Ultraviolet Photodissociation Mass Spectrometry

4.1 Overview

The step-wise reduction of dihydrofolate to tetrahydrofolate entails significant conformational changes of dihydrofolate reductase (DHFR). Binary and ternary complexes of DHFR containing co-factor NADPH, inhibitor methotrexate (MTX), or both NADPH and MTX were characterized by 193 nm ultraviolet photodissociation (UVPD) mass spectrometry. UVPD yielded over 80% sequence coverage of DHFR and resulted in production of fragment ions that revealed the interactions between DHFR and each ligand. UVPD of the binary DHFR●NADPH and DHFR●MTX complexes led to an unprecedented number of fragment ions containing either an N-terminal or C-terminal protein fragment still bound to the ligand via retention of non-covalent interactions. In addition, holo-fragments retaining both ligands were observed upon UVPD of the ternary DHFR●NADPH●MTX complex. The combination of extensive holo and apo fragment ions allowed the locations of the NADPH and MTX ligands to be mapped, with NADPH associated with the adenosine binding domain of DHFR and MTX interacting with the B loop. These findings are consistent results with previous crystallographic evidence. Comparison of the backbone cleavage propensities for apo DHFR and its holo counterparts revealed significant variations in UVPD fragmentation in the regions expected to experience conformational changes upon binding NADPH, MTX, or both ligands. In particular, the subdomain rotation and loop movements which are believed to occur upon formation of the transition state of the ternary complex are reflected in the UVPD mass spectra. The UVPD spectra indicate enhanced backbone cleavages in

regions that become more flexible or as suppressed backbone cleavages for those regions either shielded by the ligand or involved in new intramolecular interactions. This study corroborates the versatility of 193 nm UVPD mass spectrometry as a sensitive technique to track enzymatic cycles that involve conformational rearrangements.

4.2 INTRODUCTION

The use of mass spectrometry in the field of structural biology has accelerated in recent years due to the advent of more effective means of transporting native-like proteins into the gas phase, as well as refinement of methods used to probe protein structures such as hydrogen-deuterium exchange and tandem mass spectrometry.[1,2] Many of the original studies that explored the use of mass spectrometry for structural biology applications utilized covalent labeling or crosslinking of proteins and protein-ligand complexes in solution, followed by proteolytic digestion, separation, identification, and sometimes quantification of probe-modified or crosslinked peptides to reveal information about the solvent accessibility and interacting regions of the proteins.[1] More elegant strategies based on hydrogen-deuterium exchange of the protein backbone amide hydrogens have gained popularity due to their single residue resolution and sensitivity to conformational changes.[2] Most of these methods have utilized a bottom-up workflow in which the proteins of interest are enzymatically digested prior to analysis of the constituent peptides, thus providing an indirect means to correlate protein structure with peptide-level outcomes. Since the development of new, widely-accessible high performance mass spectrometers, there has been growing interest in employing top-down approaches for structural biology investigations, thus allowing evaluation of intact proteins transported to the gas phase.[3–25] Collisional, electron-based and photon-based activation methods have been used to analyze the intact proteins.[22,26–32] These methods

have proven effective for quantifying covalent labeling or hydrogen-deuterium exchange of proteins to the same degree, if not more completely, than the corresponding bottom-up approaches on proteins below 30 kDa.[3–6,8,14]

Compelling new advances in the application of mass spectrometry to structural biology problems have been inspired by the recent successes in transporting presumed native-like proteins and protein complexes to the gas phase from buffered solutions via native spray methods.[11,33] This has opened up the possibility of using mass spectrometry to directly interrogate native-like protein structures, primarily using MS/MS methods to disassemble the complexes, sequence the proteins, and draw conclusions about protein conformation based on fragmentation behavior.[10,11,15,21,22,34–39] One hallmark of native spray is the adoption of much lower charge states of the native-like proteins than those observed for denatured proteins sprayed from acidic, methanol-containing solutions. Although still a subject of debate, it is believed that proteins retain native-like conformations to a large extent, thus allowing examination binding interactions of protein complexes via an array of MS/MS methods.[11,20,21,40] Ion mobility mass spectrometry (IMMS) also plays a pivotal role in discerning and studying the three dimensional shapes of proteins and protein complexes in the gas phase.[16,41] Collision-based activation of the resulting protein-ligand or protein-protein complexes leads to ejection of the ligands or intact protein monomers or multimers (from multimeric protein complexes), as well as some sequence ions from the proteins.[12,15,30] Activation of native-like proteins without any bound ligands leads to formation of sequence ions that have been correlated with the B-factors of the proteins and thus the local stabilities of specific regions.[15,34,42] Ultraviolet photodissociation (UVPD) in particular provides unsurpassed levels of sequence coverage for intact proteins (denatured ones)[26,27,43] and has more recently exhibited similar levels of performance for native-like proteins and protein complexes.[15,34] For the

latter, UVPD resulted in conventional sequence ions for which the abundances of ions produced upon cleavage of inter-residue bonds mirrored the B-factors of the protein. In addition UVPD generated product ions comprised of a portion of the protein (*a,b,c, x,y,* or *z* ion) still bound through non-covalent interactions to the ligand.[15,44] One recent study showed that the abundance of the fragment ions from myoglobin varied as a function of heme binding (apo and holo states) in a way that aligned with expected conformational changes upon ligand binding.[34] Additionally it was recently reported that conformers of ubiquitin separated in the gas phase by ion mobility were differentiated based on their UVPD fragmentation patterns.[39] Electron capture dissociation, electron transfer dissociation, and surface induced dissociation have also been used to disassemble native proteins or protein complexes in the gas phase.[20,22,36,45–48] Collectively, these studies have provided growing evidence that native spray, in conjunction with tandem mass spectrometry, can be employed to address increasingly advanced questions about the nature of protein-ligand interactions in the context of structural biology.[13,49] Herein we report the use of UVPD to examine binary and ternary protein ligand complexes, as described for dihydrofolate reductase (DHFR).

DHFR is a small protein (~19 kDa, 186 amino acids) that reduces dihydrofolate (DHF) to tetrahydrofolate (THF). THF is the precursor for all folate coenzymes involved in numerous biosynthetic pathways.[50,51] Nicotinamide adenine dinucleotide phosphate (NADPH) serves as a co-factor of DHFR during its catalytic cycle and is oxidized in the process.[52] The step-wise DHFR reaction has been explored by many methods to elucidate the energetics and structural intermediates.[51,53–57] An outstanding depiction of the dynamic landscape of DHFR catalysis and structural models of key intermediates in the catalytic cycle are provided in ref. 56. There are two well defined binding regions in DHFR, the adenosine-binding domain and the loop-binding domain.[51] Upon both substrate and co-

factor binding, it has been established that there are specific loop movements, particularly in the M20, F-G and G-H loops, responsible for catalyzing the reduction of DHF and releasing THF and the oxidized NADP+.[51,53,57–59] These loop movements have been shown to be key conformational changes through the catalytic cycle. Additionally, DHFR plays a pivotal role in the folate cycle that produces thymine.[60] Thymine is required for the proliferation of rapidly dividing cells, such as those during cancerous metastasis. This feature makes DHFR a compelling drug target for the development of clinical therapeutics, such as methotrexate (MTX).[51,53,61,62] This potent inhibitor of DHFR has also been used in combination with NADPH to aide in the creation of models of the transition state of the DHFR reaction.[51] The fact that DHFR binds both NADPH and MTX makes it an attractive candidate for expanding the scope of UVPD for investigation of protein-ligand complexes. We report here that UVPD can be used to probe DHFR as well as its binary and ternary complexes (DHFR•NADPH, DHFR•MTX, and DHFR•NADPH•MTX) and to explore the stepwise impact of loop movements and the ligand binding sites as revealed by the UVPD fragmentation trends.

4.3 EXPERIMENTAL

4.3.1 DHFR Production

For expression of DHFR, the *E. coli folA* gene encoding dihydrofolate reductase including a C-terminal His6-tag (amplified from DH10B genomic DNA) was cloned into the NdeI/PacI sites of pETDuet-1 (Novagen). BL21(DE3) cells containing pETDuet-DHFR were cultured in 2 L of LB medium and induced with IPTG during mid log phase. Cells were harvested by centrifugation at 8000 x *g* for 10 min and resuspended in 20 mL of wash buffer (100 mM Tris, 150 mM NaCl, 1 mM EDTA at pH 8.0) with protease inhibitor cocktail (cOmplete, mini EDTA free, Roche) and lysozyme at 1 $mg.mL^{-1}$.

Following a 20 min incubation at 4 °C cells were lysed by sonication (Model 500, Fisher Scientific) and clarified three times by centrifugation at 35000 x *g* for 30 min. Lysate was filtered through a 0.2 μm membrane and DHFR was recovered by IMAC using Ni-NTA resin and gravity flow columns. Eluate was concentrated and dialyzed against 50 mM NH_4OAc pH 6.5, followed by purification to apparent homogeneity by size exclusion fast protein liquid chromatography (FPLC).

4.3.2 Mass Spectrometry

DHFR solutions were prepared by incubation of 15 μM DHFR stock in 50 mM ammonium acetate at pH 6.5 with 20x excess NADPH, methotrexate (MTX), or NADPH+MTX, respectively, for 30 minutes at room temperature. All ligands and reagents were obtained from Sigma Aldrich. Each sample was cleaned six times with a 10 kDa molecular weight cut-off filter (Millipore, Darmstadt, Germany) and diluted to 7 μM with 50 mM ammonium acetate pH 6.5 for the native protein samples, while the denatured protein sample was filtered into 50:49:1 (water:acetonitrile:formic acid). Samples were infused using a 40 nm Au-coated static tip electrospray set-up with an applied voltage of 1.2-1.5 kV. A heated capillary was set at 200°C to assist desolvation of the proteins. All experiments were undertaken on a Thermo Scientific Instruments Orbitrap Elite mass spectrometer (Bremen, Germany) equipped with a Coherent Excistar excimer laser (Santa Cruz, CA) for photodissociation.[27] UVPD was performed in the HCD cell operated at 10 mTorr helium pressure and using a single 3.0 mJ laser pulse (193 nm wavelength) per spectrum. The 9+ charge state species was selected in every case using an isolation width of 15 m/z and an AGC target of 1e5 with a maximum injection time of 1 s. 250 scans total were averaged for each spectrum.

4.3.3 Data Analysis

UVPD mass spectra were deconvoluted using the Thermo Xtract algorithm with a S/N ratio of 3, then searched through Prosight PC 3.0 modified with custom code to include nine UVPD-type ions (*a, a•, b, c, x, x•, y, y-1, z*). These ion types are consistently found by UVPD.[27] The searches for and assignments of ligand-containing fragment ions (holo fragments in which a segment of the protein retains the ligand via non-covalent interactions) were done in a manner to account for mass shifts corresponding to retention of the MTX, NADPH, or MTX+NADPH ligands by the standard sequence ion types (a,b,c,x,y,z). The ligand mass shifts plus up to three extra hydrogens (due to the prominence of hydrogen atom (not proton) migrations during activation of proteins by UVPD) were treated as variable modifications and were applied as follows: mass shift of 454.1713-457.1948 Da for MTX, 743.0750-746.09845 Da for NADPH, and 1198.2541–1201.2776 Da for MTX+NADPH. All identified ions were normalized relative to the total ion current of the respective spectra to allow evaluation of trends for all spectra. Fragment ions arising from cleavage of the backbone positions between pairs of adjacent amino acids in the protein sequence were collectively summed. For example, all product ions arising from backbone cleavages that occur N-terminal to a specific amino acid (yielding a_n, b_n, and c_n ions) were summed with all the complementary C-terminal product ions arising from the next flanking and specific amino acid containing backbone cleavage site (yielding complementary $x_{R-(n+1)}$, $y_{R-(n+1)}$, and $z_{R-n(n+1)}$ ions) where R is the total number of amino acids in the protein. For visualization of the results, residues that displayed a significant change in fragmentation yields were highlighted on crystal structures representing DHFR•NADPH, DHFR•MTX and DHFR•NADPH•MTX complexes. The PDB codes for these structures are 1RX1, 1RG7 and 1RX3, respectively. B-factors were extracted from the respective PDB files, as well as for the apo-DHFR form from 5DFR.

All visualizations were performed using Pymol 1.3 software. Pymol was also used to search for polar contacts with a 4.0 Å cutoff. All UVPD experiments were repeated at least three times. A structural representation of DHFR based on X-ray crystal structure 1RX3 is shown in **Figure 1** with the helices and binding domains labelled and color-coded.

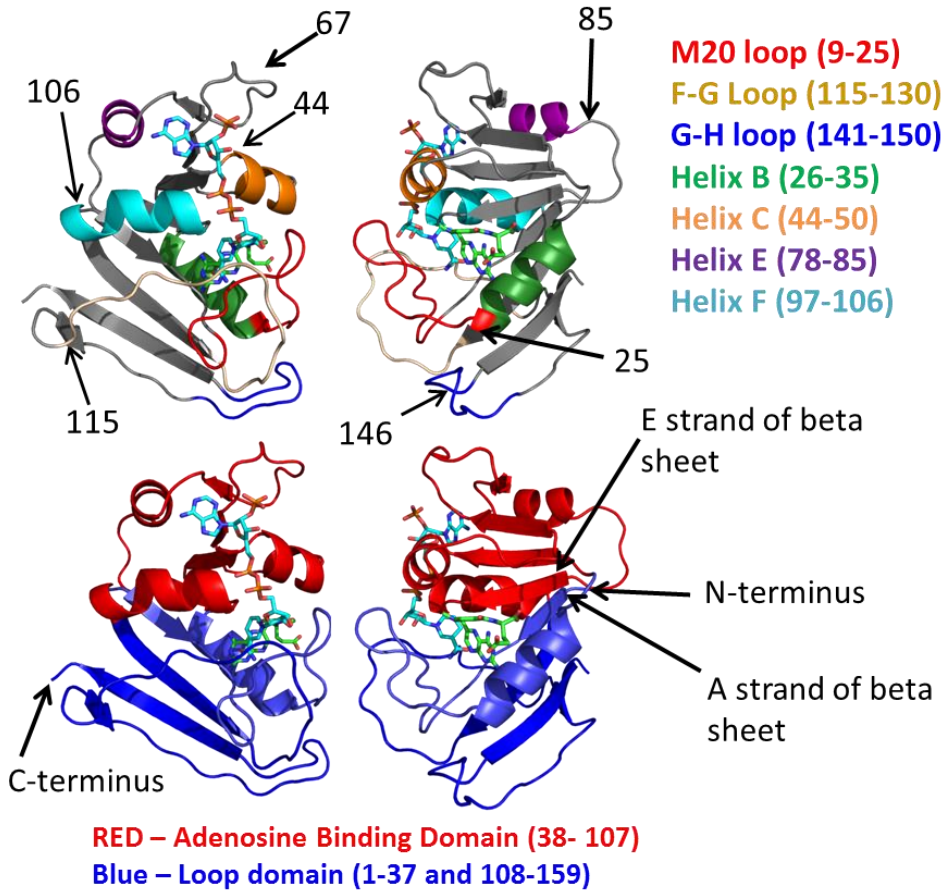

Figure 1. Schematic figure of DHFR highlighting the subdomains as well as the well-defined loop regions. Selected residues are numbered. NADPH is seen in blue sticks and MTX is shown in green sticks. This structure was sourced from 1RX3.

4.4 RESULTS AND DISCUSSION

In this study, UVPD was used as a tool to characterize previously known ligand binding sites and evaluate changes in the structure of DHFR based are reflected by variations in the fragmentation patterns of the protein. DHFR was successfully transferred to the gas phase via native ESI conditions to generate apo and holo non-covalent complexes containing methotrexate (MTX), NADPH, or both NADPH and MTX in low charge states (8+, 9+) characteristic of native-like proteins (**Figure 2**).

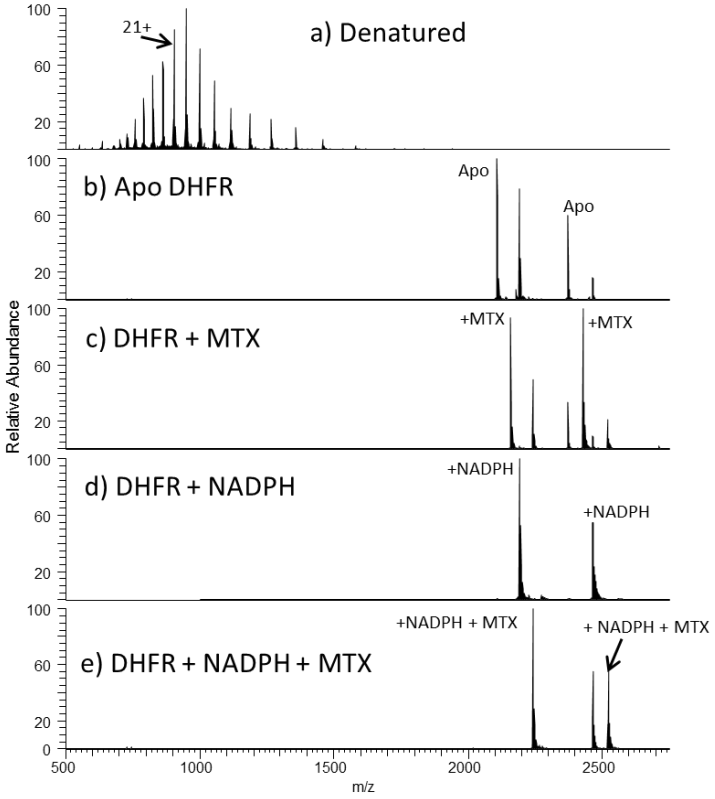

Figure 2. ESI mass spectra of a) denatured DHFR, b) native apo-DHFR, c) native DHFR + MTX, d) native DHFR + NADPH, and e) native DHFR + NADPH + MTX solutions. The DHFR stock solution contained NADPH from purification which explains the presence of DHFR●NADPH complexes in the DHFR spectrum.

The 9+ charge state was the most abundant for each DHFR complex, and thus this charge state was isolated and subjected to UVPD for all MS/MS experiments. Examples of the resulting UVPD mass spectra are displayed in **Figure 3**.

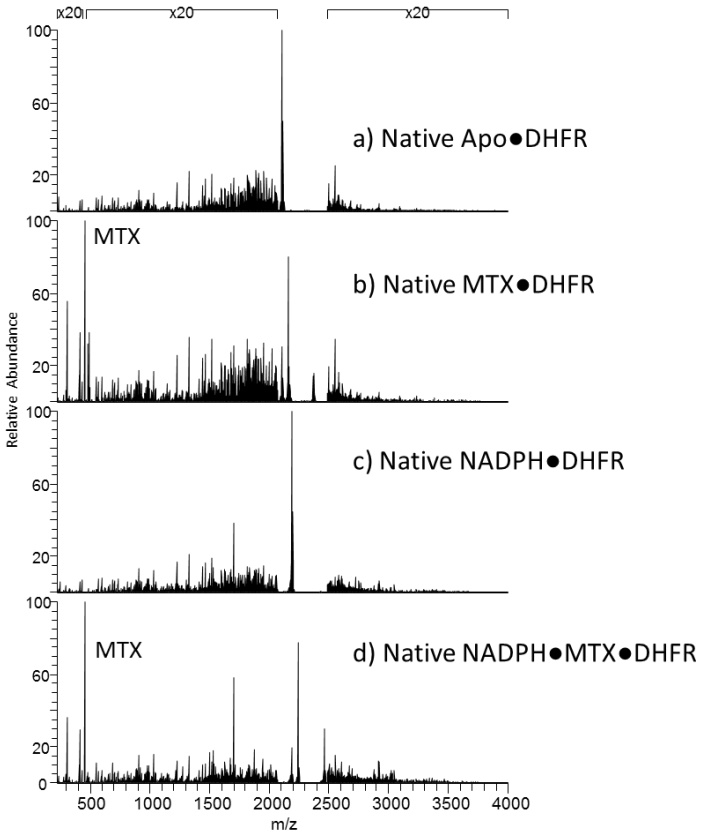

Figure 3. UVPD mass spectra of the 9+ species of a) native apo-DHFR, b) native (DHFR●MTX), c) native (DHFR●NADPH), and d) native (DHFR●NADPH●MTX). UVPD was performed using a single 193 nm 3 mJ pulse.

Deconvolution shows that the UVPD spectra are very rich and high quality, containing a variety of multi-charged *a, b, c, x, y,* and *z* ions, some retaining the NADPH and/or MTX ligands (**Figures 3** and **4**). In general *a/x*-type ions were about twice as abundant as *b/y* and *c/z* ions. UVPD of each of the native apo and holo proteins (9+) yielded at least 81%

sequence coverage. For the native-like DHFR complexes, the pattern of backbone cleavages of DHFR (based on formation of N-terminal *a, b,* and *c* ions and C-terminal *x,y*, and *z* ions by UVPD) is influenced by the presence or absence of bound ligands, as well as conformational changes that alter the flexibility/stability of specific regions and thus the susceptibility to fragmentation. The fact that absorption cross-sections of proteins may vary with charge density or conformation has been postulated previously[63] and echoed in one of our previous studies.[34] Variations in molar absorptivities of structural elements of proteins in solution (alpha-helices versus coiled/loop regions) have similarly been documented.[64]

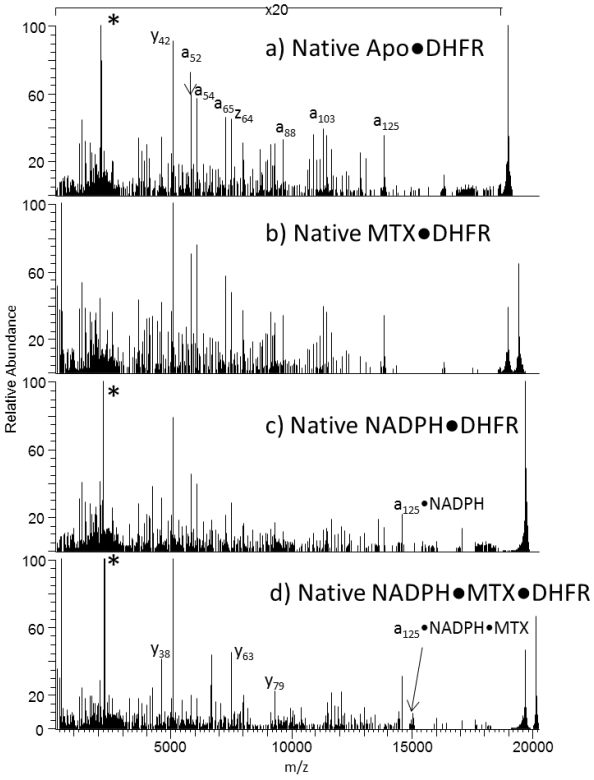

Figure 4. Deconvoluted UVPD mass spectra of the 9+ species of a) native apo-DHFR, b) native (DHFR●MTX), c) native (DHFR●NADPH), and d) native (DHFR●NADPH●MTX). UVPD was performed using a single 193 nm 3 mJ pulse.

The complexity of the UVPD mass spectra required careful analysis to differentiate and assign ligand-free (apo) and ligand-containing (holo) sequence ions, as described in the experimental section. For the ternary DHFR●NADPH●MTX complex, fragment ions containing either NAPDH or MTX or both NADPH and MTX were identified. The average numbers of ligand-containing fragment ions found for each DHFR complex are summarized in **Figure 5**.

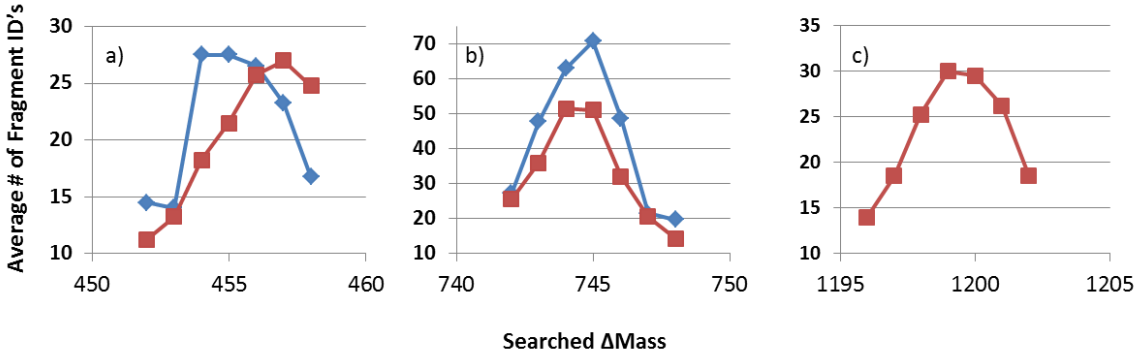

Figure 5. The average number of holo fragment ions identified containing a) MTX, b) NADPH, or c) MTX●NADPH produced upon UVPD of the binary complexes (DHFR●MTX) (in blue in a) or (DHFR●NADPH) (in blue in b) or the ternary (DHFR●NADPH●MTX) complex (red in a,b, and c).

The DHFR●MTX complex produced 46 unique methotrexate-containing fragment ions, and 135 unique NADPH-containing fragment ions were obtained from the DHFR●NADPH complex. For the multi-ligand ternary DHFR●NADPH●MTX complex, a total of 212 unique holo fragments were identified upon UVPD, including 106 retaining NADPH, 36 retaining methotrexate, and 70 containing both NADPH and methotrexate. In summary, NADPH was retained in a greater number of sequence ions than was MTX, and MTX was retained in more sequence ions for the ternary DHFR●NADPH●MTX complex than for the binary DHFR●MTX complex. In addition, NADPH was retained more frequently than MTX upon dissociation of the ternary

DHFR●NADPH●MTX complex. In fact MTX was retained more often in conjunction with retention of NADPH than retention of MTX alone. This latter result suggested that the ternary DHFR●NADPH●MTX complex engaged in different and/or stronger interactions with the MTX ligand than found in the binary DHFR•MTX complex. The detection of sequence ions retaining both NADPH and MTX for the ternary DHFR●NADPH●MTX complex was particularly interesting, thus signaling the survival of the non-covalent interactions between the protein and two different ligands during the photoactivation and dissociation process.

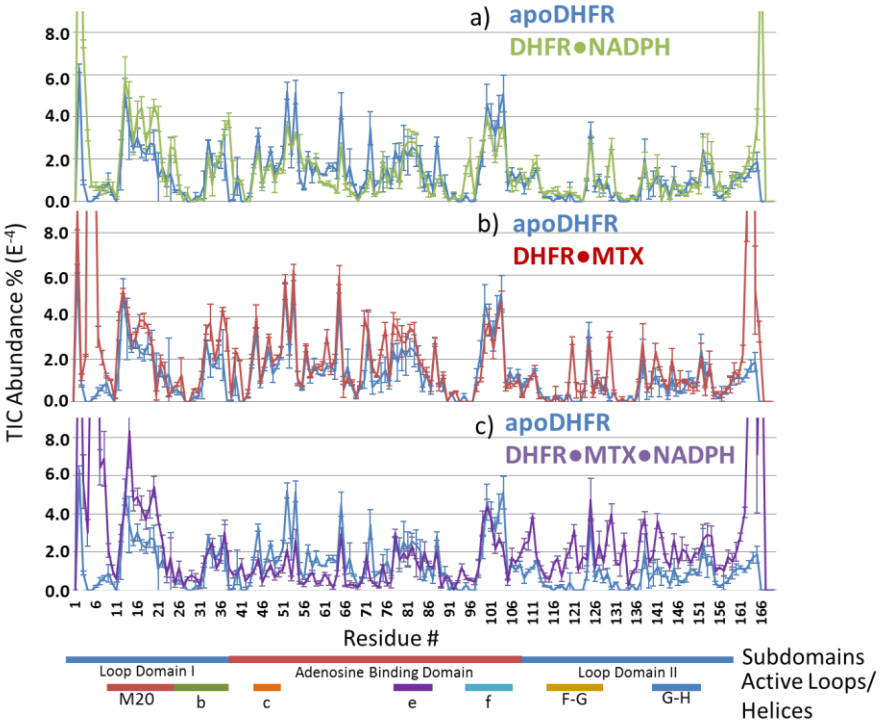

Figure 6. TIC abundance % per residue plots of summed holo + apo product ions from DHFR and its respective complexes DHFR●NADPH, DHFR●MTX, and DHFR●NADPH●MTX. The color code used for each protein is shown in the legend. Standard deviations were calculated from four replicates.

Each UVPD mass spectrum was processed to calculate the relative fragmentation propensities between each pair of amino acids in the sequence (i.e. based on backbone cleavages between every pair of residues; summing those C-terminal *x,y,z* and N-terminal *a,b,c* ions arising from each backbone site as described in the experimental section). The backbone cleavage propensities are summarized in graphical form in **Figure 6**, with the results for apo-DHFR overlaid with the corresponding results for each of the three complexes (DHFR●NADPH in **6A**, DHFR●MTX in **6B**, and DHFR●NADPH●MTX in **6C**). A number of notable differences were observed upon inspection of the fragmentation trends for apo-DHFR relative to its complexes: cleavage at some backbone sites was enhanced; others were significantly suppressed upon ligand binding. To aid in visualization of the regions of the protein for which fragmentation was enhanced or suppressed, the residues corresponding to those backbone cleavage sites are highlighted on the structures of the various DHFR complexes in **Figure 7**. The structural elements and sub-domains of DHFR are illustrated in **Figure 1**. Selected amino acids are numbered to facilitate visual orientation of the protein relative to the backbone sites for which cleavage is altered going from the apo-protein to the holo-protein complexes. In **Figure 7**, the change in fragmentation propensity for each DHFR complex is scaled relative to the fragmentation propensity of apo-DHFR. A decrease in backbone fragmentation (*suppression* of fragmentation) is highlighted in binned cool colors (cyan, blue, purple) on each crystal structure, and increases in backbone fragmentation (enhancement of fragmentation upon ligand binding) are binned into warm colors (yellow, orange, red). For native-like proteins or protein-ligand complexes in the gas phase, the suppression or enhancement of specific backbone cleavage sites upon UVPD is modulated by two primary effects: "shielding" of backbone sites by non-covalent interactions with the ligands or involvement in other intramolecular interactions (thus

suppressing fragmentation of those regions) or enhancement of backbone fragmentation due to increased conformational flexibility (higher B value).[15,34] Regions with greater conformational flexibility should have fewer intramolecular interactions than the more rigid regions. Therefore, fewer intramolecular interactions must be disrupted in order to release pairs of fragment ions (N-terminal and C-terminal products) when a particular backbone bond is cleaved, thus accounting for higher UVPD yields. Moreover, as mentioned above absorption cross-sections of different structural elements of proteins are known to vary in solution,[64] thus potentially translating to variations in local absorptivities in the gas phase.

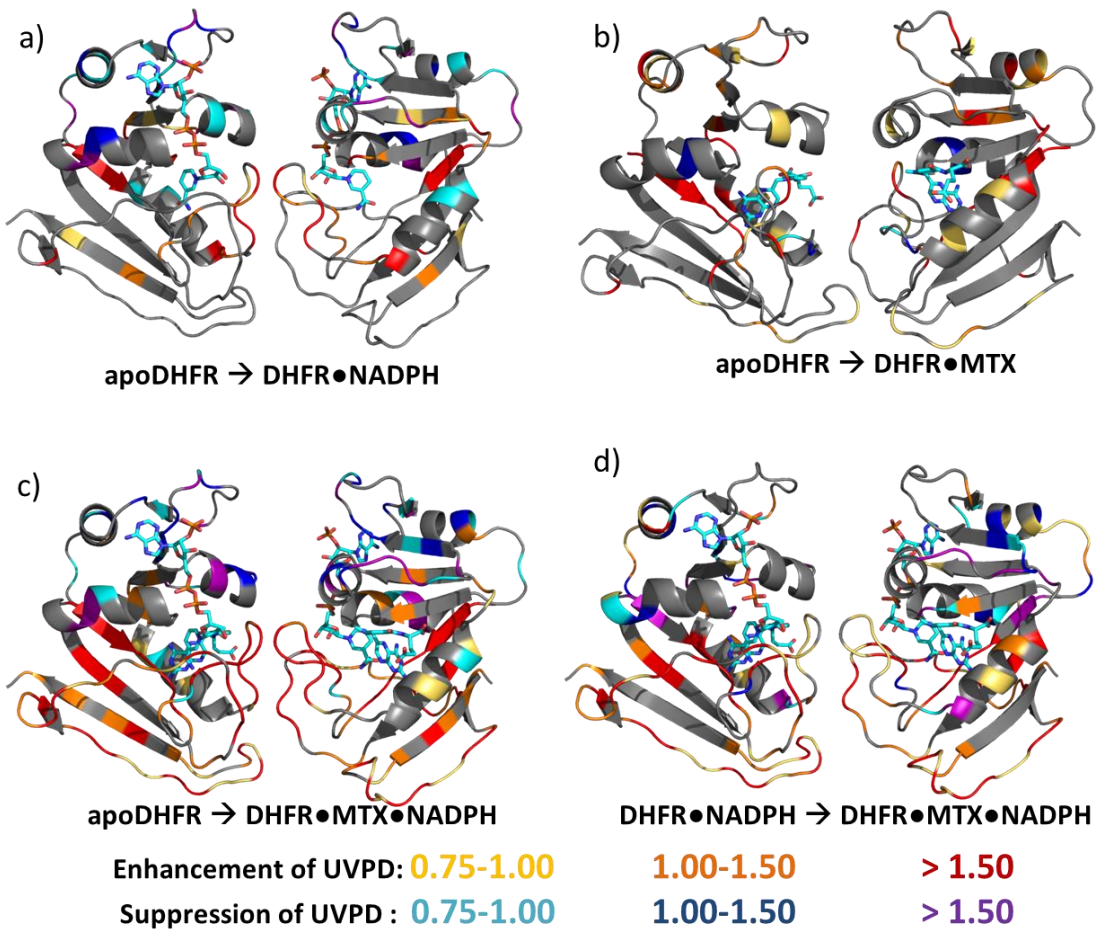

Figure 7. Those residues which had an enhanced (positive) or suppressed (negative) change in UVPD fragmentation upon comparison of apo-DHFR and a) DHFR●NADPH, b) DHFR●MTX, and c) DHFR●MTX●NADPH as well as comparison of d) DHFR●NADPH to DHFR●MTX●NADPH were highlighted according to the colored bins. In each case, the increase (enhancement) or decrease (suppression) in UVPD fragmentation yield is shown as a percentage representing the change in TIC abundance. Crystal structures 1RG7, 1RX1, and 1RX3 were used to represent the DHFR● MTX, DHFR●NADPH, and DHFR●MTX●NADPH complexes, respectively.

Upon comparison of the backbone fragmentation trends for apo-DHFR and DHFR●NADPH, the most notable enhancement in fragmentation occurs in the region of

the M20 loop (particularly backbone cleavages associated with residues 12-25), as well as from residues 37 to 41 (**Figure 6a** and **Figure 7a**). For the same comparison, significant suppression of UVPD fragmentation occurred in the adenosine binding domain (ABD) in the loop regions from residues 51 to 54, 62 to 65, and 84 to 88. There were only a few backbone cleavages in the ABD that were enhanced for the DHFR●NADPH complex relative to DHFR (cleavages adjacent to residues 55, 56, 95, and 96). Some suppression of UVPD occurred in the helices of the ABD (namely in the C, E, and F helices), most notably adjacent to residues 102-104 in the F-helix. This suggests that NADPH binding causes a conformational shift of the M20 loop as well as in the ABD, specifically in the loop region containing residues 62 to 65. These changes in the UVPD pattern parallel some of the insights drawn by Sawaya *et al.* whom reported the most in-depth crystallographic analysis of the loop and subdomain movements of DHFR and its NADPH complexes.[51] They mapped the binding of the adenosine portion of NADPH to the ABD as well as the shift the M20 loop from a disordered or open conformation to a closed conformation upon NADPH binding.[51] Sawaya *et al.* also reported that the transition from an open to closed state resulted in disruption of hydrogen bonds between residues N23 and S148.[51] While a large change in the UVPD cleavage pattern is not observed next to residue S148, there is a substantial enhancement in backbone cleavage adjacent to residues L24 and P25 which directly neighbor N23. Also notable is the lack of significant changes in cleavage of the backbone sites in the F-G loop which are presumed to maintain hydrogen bonds with the M20 loop in the DHFR●NADPH complex.[51]

Sawaya *et al.* determined that apo-DHFR and its DHFR●MTX complex had similar crystal structures overall; however, the M20 loop was disordered in the apo-DHFR structure but upon binding MTX the M20 loop adopted a more stable open

conformation.[51] The main interaction between DHFR and MTX was purported to involve the (p-aminobenzoyl)glutamate (pABG) section of MTX and helix B of DHFR (see structural guide in **Figure 1**) which arose from significant hydrophobic interactions.[51] Such interactions are anticipated to be diminished in the gas phase, thus rationalizing both the lower retention of MTX in the fragment ions produced upon UVPD of the binary DHFR●MTX and ternary DHFR●MTX●NADPH complexes as well as the relatively modest amount of change in fragmentation between the apo-DHFR and DHFR●MTX complexes (**Figure 6b**). However, there are still some notable enhancements in UVPD fragmentation for the DHFR●MTX complex relative to apo-DHFR. These changes are most prominent in the M20 loop, specifically enhancements in backbone cleavage adjacent to residues 17-18 and in helix B (next to residues 34,37,38,40, and 41). These changes in UVPD within the M20 loop are rationalized based on an ordering of the loop upon MTX binding and the interaction of helix B with MTX. In comparison to the DHFR•NADPH complex, there is little suppression of backbone cleavage of DHFR in the DHFR●MTX complex and virtually no changes in UVPD in the ABD region (which is remote from the MTX binding site based on the crystal structure).

Upon inspection of the UVPD fragmentation trends for the ternary DHFR●MTX●NADPH complex (**Figure 6c** and **Figures 7c** and **d)**, there is a significant increase in the number of backbone cleavage sites suppressed or enhanced for the ternary complex in comparison to apo-DHFR or the binary complexes. Most of the changes mirror those already noted upon the binding of NADPH, such as suppression of fragmentation in the ABD and enhancement of fragmentation in the M20 loop. However, there are new enhancements when both ligands are affiliated with DHFR. These new regions of enhancement occur in the F-G and G-H loops and in the central beta sheet of the enzyme.

It has been suggested previously that the ternary DHFR●MTX●NADPH complex mimics and exhibits properties similar to that of the active protein transition state for the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF).[51] Upon reduction of DHF, the M20 loop adopts an occluded state which results in disruption of hydrogen bonds between residues G121 and D122 from the F-G loop and residues G15 and M16 from the M20 loop.[51] Additionally upon this M20 loop transition, residue S148 from the G-H loop reforms a hydrogen bond to N23.[51] Other conformational changes occur during this transition-like state, including binding of NADPH and a twisting of the beta-sheet between strands A and E. These conformational changes are mirrored by large enhancement of backbone fragmentation upon UVPD, spanning both the F-G loop (next to residues 115,116, 119-123, 125, 127-130) and G-H loop (next to residues 141-150), and with a lower degree of enhancement of fragmentation in the M20 loop (residues 14-16, 18-20, and 22). This low enhancement of UVPD in the M20 loop suggests that the M20 loop may be much less dynamic and instead maintains a more closed and rigid conformation. A small amount of suppression of fragmentation is also observed adjacent to residues 24 and 25 in the M20 loop. Interestingly, as noted above for the binary DHFR●NADPH complex the cleavage adjacent to these residues (24 and 25) was enhanced, presumably because the hydrogen bond between S148 and N23 was disrupted upon NADPH binding. All of these changes from UVPD tracked well with the transition state mechanism of DHFR discussed above. While that transition state explanation evolved from a crystallographic analysis of solid state (static) structures,[50] the UVPD trends offer a dynamic depiction dependent on disruption and formation of hydrogen bonds or other electrostatic interactions that influence the propensity for backbone cleavages upon UVPD. There are also compelling changes in UVPD fragmentation occurring in the region of DHFR that directly contacts the MTX ligand. Certain backbone

cleavages increased in helix B (adjacent to residues 29, 32, and 34), whereas there was a significant suppression of backbone cleavages in the ABD region nearest the MTX binding site (next to residues 52,53 and 55-59). This is suggestive of a stronger binding interaction occurring in the ternary DHFR●MTX●NADPH complex in comparison to the DHFR●MTX binary complex discussed above.

The pattern and relative abundance of ligand-containing (holo) sequence ions produced upon UVPD are displayed as a function of the backbone cleavage site in **Figure 8** and are classified as N-terminal or C-terminal ions. These types of maps are useful for discerning the location of the ligands upon dissociation of the non-covalent protein-ligand complexes.

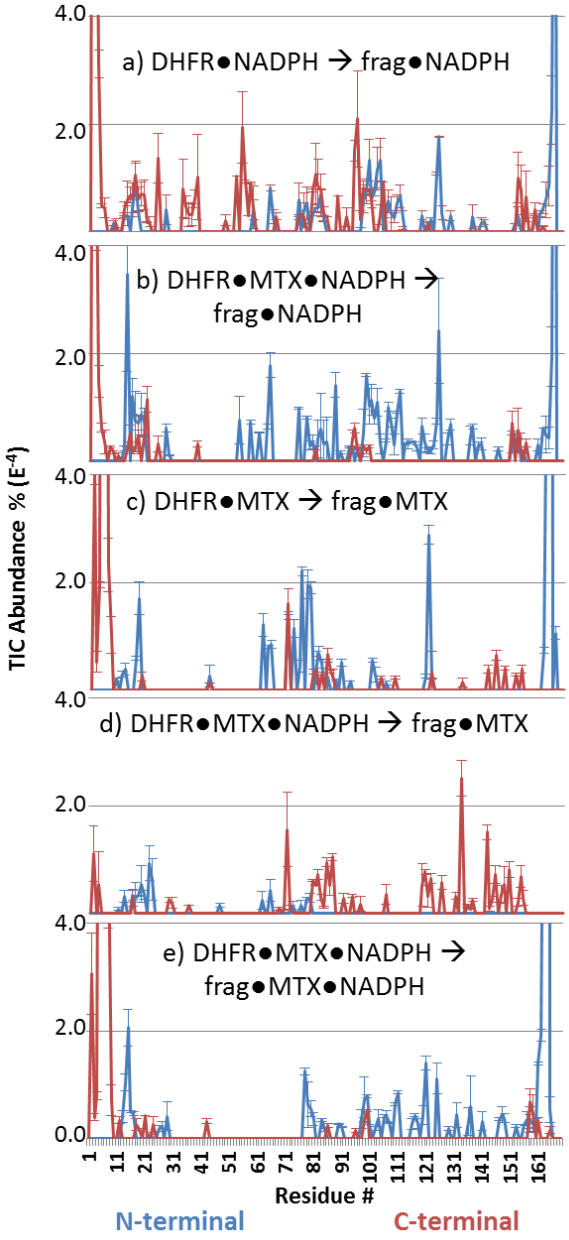

Figure 8. Plots of the N-terminal (blue) and C-terminal (red) holo fragment ions shown relative to the backbone cleavage site and produced upon UVPD of DHFR•NADPH, DHFR•MTX and DHFR•NADPH•MTX complexes, including those fragment ions retaining MTX arising from (a) DHFR•MTX and (b) DHFR•NADPH•MTX, those fragment ions retaining NADPH arising from (c) DHFR•NADPH and (d) DHFR•NADPH•MTX, and those fragment ions retaining both MTX and NADPH arising from (e) DHFR•NADPH•MTX.

The identified holo ions represent a significant portion of the fragment abundance per residue and sometimes are responsible for 100% of the ion current affiliated with backbone cleavages at a particular residue (**Figure 9**).

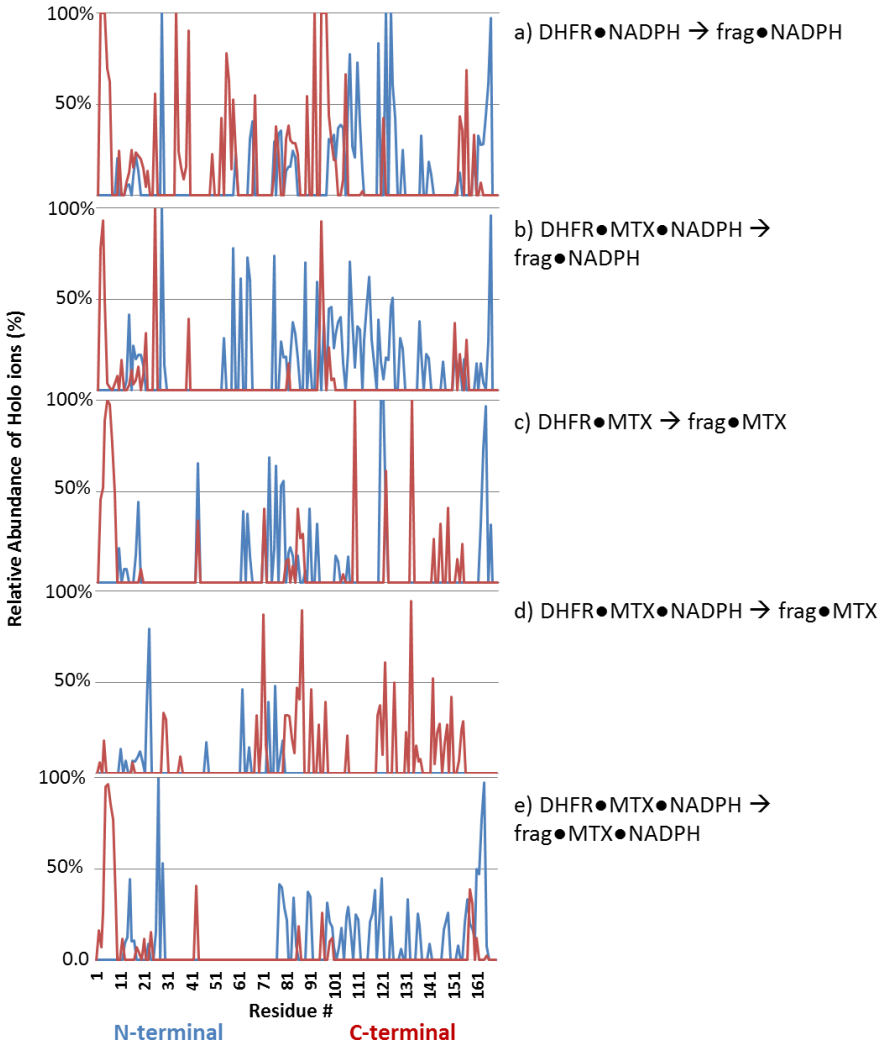

Figure 9. Plots of the relative abundance of N-terminal (blue) and C-terminal (red) holo fragment ions versus the total abundance of fragment ions at each residue. This data was produced by UVPD of DHFR•NADPH, DHFR•MTX and DHFR•NADPH•MTX complexes, including those fragment ions retaining MTX arising from (a) DHFR•MTX and (b) DHFR•NADPH•MTX, those fragment ions retaining NADPH arising from (c) DHFR•NADPH and (d) DHFR•NADPH•MTX, and those fragment ions retaining both MTX and NADPH arising from (e) DHFR•NADPH•MTX.

As expected, those fragment ions that retain NADPH for the DHFR●NADPH complex differ from those that retain MTX for the DHFR●MTX complex, an outcome anticipated based on the different binding sites of each ligand. For example, the NADPH ligand is known to interact with residues in the M20 loop as well as those found in the ABD based on the extensive array of N- and C-terminal holo fragment ions that contain these same stretches of amino acids. These amino acids are highlighted in purple in the space-filled model of DHFR in **Figure 10** and include residues 13, 14, 16-18, 59, 75, 77, 80-84, 98-101, 104, 112, 153, 159, and 162. This result is generally consistent with the interaction of NADPH with the M20 loop (residues 9-25) and the ABD (residues 38-107),[50] in particular involving polar contacts between DHFR (residues 7, 18, 19, 44-46, 63, 64, 96-98, and 102) and NADPH. (Polar contacts were found using PyMol with a 4 angstrom cutoff from analysis of crystal structure 1RX1.) The holo fragment ions produced upon UVPD do not overlay perfectly with the contact residues 7, 44-46 or 63-64 predicted from PyMol. However, there are both N-terminal and C-terminal holo-ions containing NADPH based on backbone cleavages adjacent to the predicted residues at 9 and 10 as well as 65-67. The proximity of these residues with those from the PyMol prediction provides compelling evidence that the holo fragment ions produced by UVPD reflect the NADPH binding site.

The MTX ligand was less consistently retained following UVPD and thus fewer informative holo sequence ions were produced (**Figure 8c**). There are relatively few overlapping C-terminal and N-terminal holo ions that are particularly diagnostic for elucidation of the ligand binding site. The MTX-containing fragment ions occurred from backbone cleavages adjacent to residues 43, 71, 81, 83, 85, and 122. There were also N- and C-terminal holo ions from backbone cleavages at residues 18 and 19. In general MTX is known to primarily interact with residues in the loop domain (1-37 and 108-159)

and specifically with residues 5, 27, 52, 57, 94, and 100 based on polar contacts from crystal structure 1RG7.[50] The low frequency of holo fragment ions prohibited more detailed assessment of the MTX binding site based on the UVPD data.

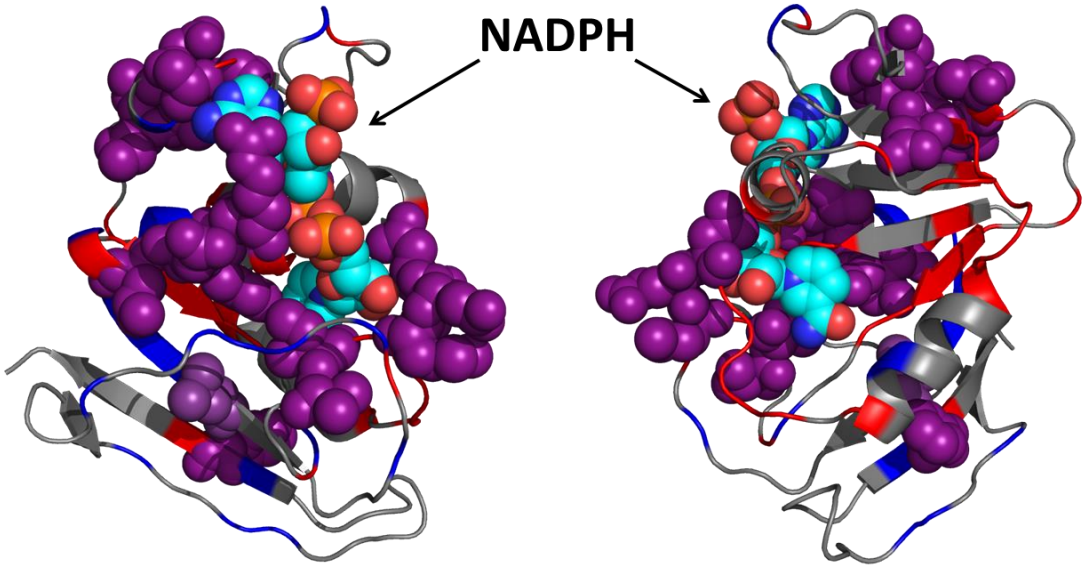

Figure 10. Space filling model of NADPH (in multicolor spheres) and the predicted (purple spheres) interacting residues from UVPD fragmentation. These residues show overlapping N and C termini holo ions. Other holo ions from the N-terminus are also highlighted in blue and others from the C-terminus are highlighted in red (non space filling).

UVPD of the ternary DHFR●NADPH●MTX complex resulted in an extensive array of NADPH-containing fragment ions (**Figure 8b** and **8e**) encompassing many of the same ones as observed for the DHFR●NADPH complex, with some exceptions. For the ternary DHFR●NADPH●MTX complex there are a more elaborate array of C-terminal ions and a stretch of missing C-terminal ions corresponding to cleavages adjacent to residues 34 to 38 that were notable for the binary DHFR●NADPH complex. The latter residues encompass or border helix B. Helix B is known to directly interact with MTX,[50] thus explaining the potential suppression of cleavage of those backbone sites upon UVPD of the ternary complex in the present study. Additionally, the

production of MTX-containing N-terminal fragments arising from backbone cleavages from residues 13 to 23 was observed upon UVPD of the ternary DHFR●NADPH●MTX complex (**Figure 8d** and **8e**) in addition to backbone cleavages from 9-12 and 17-19 previously noted upon UVPD of the binary DHFR•MTX complex. The enhancement in the number and variety of MTX-containing fragment ions for the ternary complex suggests some sort of modulation of MTX binding in conjunction with NADPH binding which causes the M20 loop (residues 9-25) to change conformation.

There is less noticeable overlap in the sequence ions that retain solely MTX or solely NADPH for the ternary DHFR●NADPH●MTX complexes relative to the corresponding holo sequence ions observed upon UVPD of the binary DHFR●NADPH and DHFR●MTX complexes (i.e. **Figure 8a** versus **8b** and **8c** versus **8d**). For example, there are many more C-terminal ions that retain NADPH for DHFR●NADPH (**Figure 8a**) than for DHFR●NADPH●MTX (**Figure 8b**). In contrast, there are more N-terminal ions that retain MTX for DHFR●MTX (**Figure 8c**) than for DHFR●NADPH●MTX (**Figure 8d**). This may be explained by the fact that NADPH is mainly associated with the ABD which is largely composed of the middle of the N-terminal half of the protein, and MTX is known to interact with the helix B region of the loop domain, thus "splitting" the holo N-terminal ion current. In terms of the fragment ions that retain *both* ligands upon UVPD of the ternary DHFR●NADPH●MTX complexes, the N-terminal fragments showed significant overlap with the N-terminal NADPH-containing fragment ions from DHFR●NADPH (**Figure 8a** versus **8e**). This results suggests that NADPH does not "migrate" significantly upon addition of the second ligand (MTX) nor does the UVPD process cause significant ligand rearrangement. The fragment ions that retained both MTX and NADPH upon UVPD of the DHFR●NADPH●MTX complexes are for the most part ones that mirror ones that retained NADPH or MTX alone. This

complementarity is consistent with maintenance of the same networks of non-covalent interactions that retain NADPH at one region of the protein and MTX at another.

4.5 CONCLUSION

This study demonstrates the versatility of 193 nm UVPD for deciphering protein-ligand interactions in native-like ternary protein complexes. The retention of non-covalent interactions upon UV photoactivation allows the determination of ligand binding sites based on observation of ligand-containing (holo) product ions as well as suppression of backbone cleavages compared to the same backbone cleavages in the ligand-free (apo) protein. The UVPD trends correlate well with previously determined crystal structure information. Interpretation of the two pools of product ions (holo and apo) gives a more complete picture of the structural changes that DHFR undergoes upon NADPH and MTX binding. In particular, the changes in abundances of the holo and apo fragment ions reflect the conformational changes of DHFR that accompany the ligand binding events. The large array of holo fragment ions containing an N-terminal or C-terminal portion of the protein and either NADPH, MTX or both NADPH and MTX ligands allowed the locations of each ligand to be determined in a manner that reflected the subdomain rotation and loop movements of DHFR upon interaction with the two ligands. The ability to examine ternary complexes (in addition to binary complexes and individual proteins) makes UVPD-MS a compelling new approach for addressing increasingly complicated questions in the arena of structural biology with numerous potential implications for studies of enzyme inhibitors and the drug discovery process. The potential sensitivity of UVPD to the impact of single point mutations on protein conformation is a related application that is currently underway.

4.6 REFERENCES

(1) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(2) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(3) Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.; Brodbelt, J. *Anal. Chem.* **2014**, *86* (5), 2534–2542.
(4) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. *J. Am. Chem. Soc.* **2009**, *131* (35), 12801–12808.
(5) Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. *J. Mass Spectrom.* **2004**, *39* (3), 322–328.
(6) Wang, G.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (15), 7293–7298.
(7) Stefanowicz, P.; Kijewska, M.; Szewczuk, Z. *Anal. Chem.* **2014**, *86* (15), 7247–7251.
(8) Bobst, C. E.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (11), 5225–5231.
(9) Nagy, K.; Redeuil, K.; Rezzi, S. *Anal. Chem.* **2009**, *81* (22), 9365–9371.
(10) Robinson, E. W.; Leib, R. D.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (10), 1470–1480.
(11) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(12) Hall, Z.; Hernández, H.; Marsh, J. A.; Teichmann, S. A.; Robinson, C. V. *Structure* **2013**, *21* (8), 1325–1337.
(13) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.
(14) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. *J. Am. Chem. Soc.* **2008**, *130* (35), 11574–11575.
(15) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(16) Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.
(17) Simmons, D. A.; Dunn, S. D.; Konermann, L. *Biochemistry (Mosc.)* **2003**, *42* (19), 5896–5905.
(18) Wang, F.; Tang, X. *Biochemistry (Mosc.)* **1996**, *35* (13), 4069–4078.
(19) Wright, P. J.; Zhang, J.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1906–1913.
(20) Zhou, M.; Jones, C. M.; Wysocki, V. H. *Anal. Chem.* **2013**, *85* (17), 8262–8267.
(21) Breuker, K.; Brüschweiler, S.; Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, *50* (4), 873–877.
(22) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(23) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2010**, *132* (25), 8602–8609.
(24) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(25) Modzel, M.; Stefanowicz, P.; Szewczuk, Z. *Rapid Commun. Mass Spectrom.* **2012**, *26* (23), 2739–2744.
(26) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(27) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

(28) Catherman, A. D.; Durbin, K. R.; Ahlf, D. R.; Early, B. P.; Fellers, R. T.; Tran, J. C.; Thomas, P. M.; Kelleher, N. L. *Mol. Cell. Proteomics* **2013**, *12* (12), 3465–3473.
(29) Skinner, O. S.; Catherman, A. D.; Early, B. P.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L. *Anal. Chem.* **2014**.
(30) Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L. *Anal. Chem.* **2013**, *85* (23), 11163–11173.
(31) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(32) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(33) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(34) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(35) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(36) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, 1–9.
(37) Zhang, H.; Cui, W.; Gross, M. L.; Blankenship, R. E. *FEBS Lett.* **2013**, *587* (8), 1012–1020.
(38) Cui, W.; Rohrs, H. W.; Gross, M. L. *Analyst* **2011**, *136* (19), 3854–3864.
(39) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308-10314.
(40) Vahidi, S.; Stocks, B. B.; Konermann, L. *Anal. Chem.* **2013**, *85* (21), 10471–10478.
(41) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(42) Zhang, H.; Cui, W.; Gross, M. L. *Int. J. Mass Spectrom.* **2013**, *354–355*, 288–291.
(43) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.
(44) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(45) Clarke, D. J.; Murray, E.; Hupp, T.; Mackay, C. L.; Langridge-Smith, P. R. R. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (8), 1432–1440.
(46) Zhang, H.; Cui, W.; Wen, J.; Blankenship, R. E.; Gross, M. L. *Anal. Chem.* **2011**, *83* (14), 5598–5606.
(47) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. *Anal. Chem.* **2014**, *86* (1), 317–320.
(48) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(49) Harvey, S. R.; Porrini, M.; Konijnenberg, A.; Clarke, D. J.; Tyler, R. C.; Langridge-Smith, P. R. R.; MacPhee, C. E.; Volkman, B. F.; Barran, P. E. *J. Phys. Chem. B* **2014**, *118* (43), 12348–12359.
(50) Huennekens, F. M. *Adv. Enzyme Regul.* **1994**, *34*, 397–419.
(51) Sawaya, M. R.; Kraut, J. *Biochemistry (Mosc.)* **1997**, *36* (3), 586–603.
(52) Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119–140.
(53) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(54) Wang, Z.; Singh, P.; Czekster, C. M.; Kohen, A.; Schramm, V. L. *J. Am. Chem. Soc.* **2014**, *136* (23), 8333–8341.

(55) Wan, Q.; Bennett, B. C.; Wilson, M. A.; Kovalevsky, A.; Langan, P.; Howell, E. E.; Dealwis, C. *Proc. Natl. Acad. Sci.* **2014**, *111* (51), 18225–18230.
(56) Liu, C. T.; Francis, K.; Layfield, J. P.; Huang, X.; Hammes-Schiffer, S.; Kohen, A.; Benkovic, S. J. *Proc. Natl. Acad. Sci.* **2014**, *111* (51), 18231–18236.
(57) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. *Acc. Chem. Res.* **2015**.
(58) Liu, C. T.; Layfield, J. P.; Stewart, R. J.; French, J. B.; Hanoian, P.; Asbury, J. B.; Hammes-Schiffer, S.; Benkovic, S. J. *J. Am. Chem. Soc.* **2014**, *136* (29), 10349–10360.
(59) Osborne, M. J.; Schnell, J.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. *Biochemistry (Mosc.)* **2001**, *40* (33), 9846–9859.
(60) Hatse, S.; De Clercq, E.; Balzarini, J. *Biochem. Pharmacol.* **1999**, *58* (4), 539–555.
(61) Rajagopalan, P. T. R.; Zhang, Z.; McCourt, L.; Dwyer, M.; Benkovic, S. J.; Hammes, G. G. *Proc. Natl. Acad. Sci.* **2002**, *99* (21), 13481–13486.
(62) Li, R.; Sirawaraporn, R.; Chitnumsub, P.; Sirawaraporn, W.; Wooden, J.; Athappilly, F.; Turley, S.; Hol, W. G. J. *J. Mol. Biol.* **2000**, *295* (2), 307–323.
(63) Brunet, C.; Antoine, R.; Dugourd, P.; Canon, F.; Giuliani, A.; Nahon, L. *J. Chem. Phys.* **2013**, *138* (6), 064301.
(64) Tsai, C. S. *Biomacromolecules: Introduction to structure, function and informatics*; John Wiley & Sons, 2007.

Chapter 5: Characterization of Trimethoprim Resistant *E. coli* Dihydrofolate Reductase Mutants by Mass Spectrometry and P21L Inhibition by Propargyl-Linked Antifolates

5.1 Overview

As antibiotic resistance continues to grow and is being recognized as a worldwide problem efforts are being concerted to understand how mechanisms of resistances are acquired or evolved so that more effective inhibitors can be generated to attack these resistant pathogens. One potentially pathogenic bacterium, *Escherichia coli*, which is primarily responsible for bladder infections, has shown significant resistance to the standard antibiotic for this infection, trimethoprim (TMP) which inhibits dihydrofolate reductase (DHFR) thus halting tetrahydrofolate production. The resistance is achieved through single point mutations in the protein. Two clinically relevant resistant mutations were studied P21L and the W30R mutations. These mutations were characterized, along with the WT DHFR, by a variety of methods to understand the kinetics, thermodynamics, and structure of protein by using Michaelis-Menten and inhibitory kinetics, size exclusion chromatography, and native-mass spectrometry with ultraviolet photodissociation, respectively. Briefly, it was found that the W30R mutation had significantly less kinetic fitness for THF production, both mutants were more susceptible to dissociation or formed less stable complexes, and main structural changes were seen in substrate binding pocket for W30R while the P21L mutant saw the most change in the M20 loop as well as the c-terminal portion containing the essential G-H functional loop. With this compilation of these studies it is suggestive that these mutations are conferring resistance through

separate mechanisms. Additionally, a novel propargyl-linked antifolate was shown to be potent against the P21L mutation.

5.2 INTRODUCTION

Growing concerns about antibiotic-resistant strains of *E.coli*, as well as numerous other pathogenic bacteria, have spurred efforts to expand the pipeline of inhibitors and better understand their interactions with protein targets.[1–3] As one example, dihydrofolate reductase (DHFR) is a protein found in all organisms and plays a key role in converting dihydrofolate into tetrahydrofolate, a process essential for the synthesis of purines and thymidylic acid.[4] The structures of DHFR in Gram-negative bacteria are distinctive from DHFR in mammalian cells, thus allowing development of inhibitors that are selectively active for bacterial DHFR. Owing to its high affinity for bacterial dihydrofolate reductase (DHFR), trimethoprim (TMP) has been one of the most widely used antibiotics for the treatment of bladder infections.[1–3,5] However, multiple strains of *E. coli* have developed resistance to TMP through chromosomal and point mutations which ultimately modulate the structure of DHFR and allow retention of function even in the presence of previously successful antibiotics.[1–3,6,7] This growing global health threat of antibiotic-resistant bacteria has motivated the search for new inhibitors for antibiotic-resistant pathogens.[8] Specifically *E. coli* has developed the ability to become resistant to trimethoprim (TMP), one of the most commonly used antibiotics for bladder infections, through several chromosomal and point mutations in the protein dihydrofolate reductase (DHFR).[1–3,6]

The ability to assess protein interactions with high sensitivity by mass spectrometry has proven to be a powerful new approach in recent years. The spectrum of mass spectrometry strategies ranges from those that utilize covalent labelling,[9] non-covalent labelling[10] and crosslinking methods[11] to decipher solvent-accessible regions of

proteins in the absence or presence of other proteins, ligands, or inhibitors, to those that use native mass spectrometry to evaluate stoichiometries of protein complexes[12–14] and to correlate fragmentation patterns with conformational variations.[15,16] Native mass spectrometry entails spraying proteins from solutions containing high concentrations of volatile salts to add the preservation of non-covalent interactions and native-like conformations of proteins as they are transferred to the gas phase. Native mass spectrometry has been applied to membrane proteins and whole virus capsids revealing biologically relevant insights that have opened new frontiers in the application of mass spectrometry in the field of structural biology.[17–20] Although, the structural resolution obtained from mass spectrometry-based methods rarely rivals that obtained from NMR and X-ray crystallography methods, the speed and low sample consumption of mass spectrometry give it some compelling advantages.[9,10,12,21–23]

Coupling native spray mass spectrometry with non-collisional activation techniques such as electron capture dissociation, electron transfer dissociation, and ultraviolet photodissociation (UVPD) offers a way to extract additional details about the structures of proteins.[15,16,24–28] These details include insight into ligand binding localization,[16,27,29,30] conformational changes,[16,31–33] and protein-protein interfacial regions.[26,34] Native-mass spectrometry can be combined with other auxiliary methods to gain activity/thermodynamic information beyond just structural insight. For example, recently a method that integrated a size-exclusion separation method with mass spectrometry, termed kinetic size exclusion chromatography (SEC), was developed to determine dissociation constants of protein-ligand complexes.[35,36] This kinetic SEC-MS technique required the use of rather harsh conditions to dissociate the protein complexes, and only the small molecule were monitored by mass spectrometry.[35] Native MS has also been combined with SEC or ion exchange chromatography to study complex protein

systems such as oligomers of BSA and protein conjugates for bio therapeutics.[37,38] We build on this prior SEC-MS work to monitor the survival of protein complexes as well as gain qualitative thermodynamic information in the form of k_{off} upon ligand dissociation through the SEC column.

Top-down UVPD-MS has been shown to be a promising method for evaluating conformational variations in proteins upon ligand binding, and we recently used this strategy for deciphering the conformational changes of DHFR and its inhibition by methotrexate (MTX).[16] Building on this previous study, we now report the use of UVPD to explore TMP-resistant DHFR constructs (P21L and W30R) to characterize their structural changes and determine how the mutations drive antibiotic-resistance. The structural mechanism of resistance to TMP by DHFR mutations in not yet clear, whether a steric hindrance disallowing tight binding of TMP or induction of a thermodynamic or kinetic shift favoring protein activity over TMP binding. This type of insight information would useful in designing new inhibitors for future therapeutic uses. Additionally, we employ size exclusion chromatography (SEC) to monitor the dissociation of protein-ligand complexes with MS detection. This method allows evaluation of the k_{off} values of DHFR-inhibitor complexes. Specifically, two TMP-resistant DHFR variants (P21L and W30R) clinically isolated from *E. coli* are the focus of the present study.[5,6,39] In addition, a new class of DHFR inhibitors, propargyl-linked antifolates (PLAs), have been shown to be potent against an array of different species of DHFR-containing bacteria, including wild-type *E. coli*.[40–42] A set of novel inhibitors is evaluated against the two TMP *E. coli* mutants with an emphasis on probing specific interactions that contribute to the inhibition of DHFR in the present study. Through the integration of mass spectrometry and kinetic data, the relationship between structure and function is bridged to give a more complete of how these mutations cause TMP-resistance.

5.3 EXPERIMENTAL

5.3.1 *E. Coli* DHFR Purification

The *E. coli* folA gene encoding DHFR including a C-terminal His6-tag (amplified from DH10B genomic DNA) was cloned into pACYCDuet-1 (Novagen). P21L and W30R mutations were introduced by QuikChange PCR. BL21(DE3) cells transformed with pACYC-DHFR were diluted 1/250 in 0.5 L of terrific broth and induced with (0.5 mM) IPTG during mid log phase. Cells were harvested by centrifugation at 8000 x g for 10 min and resuspended in 25 mL of wash buffer (50 mM K 2 HPO 4 , 300 mM NaCl and 10% glycerol at pH 8.0) with protease inhibitor cocktail (cOmplete, mini EDTA free, Roche) and lysozyme at 0.5 mg/mL. Following 20 min incubation at 4 °C, cells were lysed by sonication (Model500, Fisher Scientific) and clarified three times by centrifugation at 35000 x g for 30 min. Lysate was filtered through a 0.2 μm membrane and DHFR was recovered by IMAC (immobilized metal ion affinity chromatogrphy) using Ni-NTA resin and gravity flow columns. Eluate was concentrated and dialyzed against 50 mM ammonium acetate (pH 6.5) followed by purification to apparent homogeneity by size exclusion FPLC.

5.3.2 Size Exclusion Chromatography-MS (SEC-MS)

All size exclusion chromatography experiments were performed using a Dionex LC system interfaced to a Thermo Scientific Instruments Velos dual linear ion trap mass spectrometer (San Jose, CA). The LC effluent was introduced using a HESI source with an applied voltage of 4 kV. A 2.1 mm x 30 cm Zenix-C column with 80 Å pore size and 3 μm particle size was used. An isocratic mobile phase comprised of 150 mM ammonium acetate at pH 6.5 was applied at a flow rate of 80 μL/min. Analytical solutions contained 12 uM protein in 150 mM ammonium acetate at pH 6.5 with 2% DMSO to facilitate the

solubilization of inhibitors. For each run 5 μg of protein was injected onto column. MS1 spectra over two *m/z* regions: 1500-4000 for proteins and complexes and a narrow low *m/z* region around the small molecule inhibitor of interest were collected in an alternating fashion. Experiments were performed in triplicate. Peak areas of the detected protein complexes were tabulated to calculate the proportion of each eluting from the column.

5.3.3 Native nano-spray Mass Spectrometry

Solutions (7 μM protein concentration) were loaded into pulled tip silica emitters coated with Au/Pd and sprayed using an applied voltage of 1.9-2.1 kV. All solutions were analyzed using a Thermo Scientific Instruments Orbitrap Elite mass spectrometer (San Jose, CA) outfitted with a 193 nm excimer laser as described earlier.[43] The buffer contained 150 mM ammonium acetate at pH 6.5 with 2% DMSO. Inhibitors were added to a 2X molar ratio relative to the protein, and NADPH was added to a 5X molar ratio. The 7+ charge state was selected for UVPD fragmentation for each protein or protein complex. Spectra were collected at 120 k resolution at 400 *m/z* with an AGC of 1e5. For UVPD, proteins were activated using a single 2.5 mJ pulse from an unfocused excimer laser. 200 scans were averaged for each spectrum. Experiments were collected and analyzed in triplicate. Thermo Xtract with a S/N of 2 was used to deconvolute each spectrum. An in-house constructed Web-application was used to analyze the resulting deconvoluted data against DHFR sequences with a 10 ppm mass error tolerance. Data was further analyzed as described in ref. 17.

5.3.4 Growth Curve Experiments

MG16550 *E. coli* cells were grown and diluted to 0.1 absorbance units at OD of 600 nm. Cells were then incubated at 37°C with 1 μg/ml or 10 μg/ml concentrations of

trimethoprim, methotrexate and the six novel PLAs. UV absorbance measurements were recorded at 600 nm for 24 hours every 5 minutes in triplicate.

5.3.5 Enzyme Kinetics

Enzyme kinetics were determined by Lineweaver−Burke plots generated by enzyme activity assays using 12.5−100 μM DHF with 20 μM NADPH for DHF K_m and V_{max} values or 12.5−100 μM NADPH with 50 μM DHF for NADPH K_m values. A double-reciprocal plot of enzyme kinetics was generated by plotting $1/V_0$ as a function of 1/[DHF] or 1/[NADPH]. The Michaelis-Menten equation was used to determine K_m and V_{max} values.

5.3.6 Enzyme Inhibition

Enzyme inhibition assays were performed by monitoring the rate of NADPH oxidation by DHFR via absorbance at 340 nm at room temperature in assay buffer containing 20 mM TES, pH 7.0, 50 mM KCl, 0.5 mM EDTA, 10 mM beta-mercaptoethanol, and 1 mg/mL BSA using 0.1 mM NADPH and 2 μg/mL enzyme. Inhibitor, in DMSO, was added to the enzyme−NADPH mix and allowed to incubate for 5 min before the addition of 0.1 mM DHF in 50 mM TES, pH 7.0. The inhibitor concentration and volume are based on the conditions that result in a 50% reduction in enzyme activity.

5.4 DISCUSSION

Upon native spray, ternary complexes containing trimethoprim and co-factor NADPH are produced for each variant of DHFR (WT, P21L and W30R) (**Figure 1**). Additionally, the MS1 spectra show that the mutation from tryptophan to arginine for variant W30R does not significantly change the resulting charge states, and the 6+ complexes are favored for all three variant.

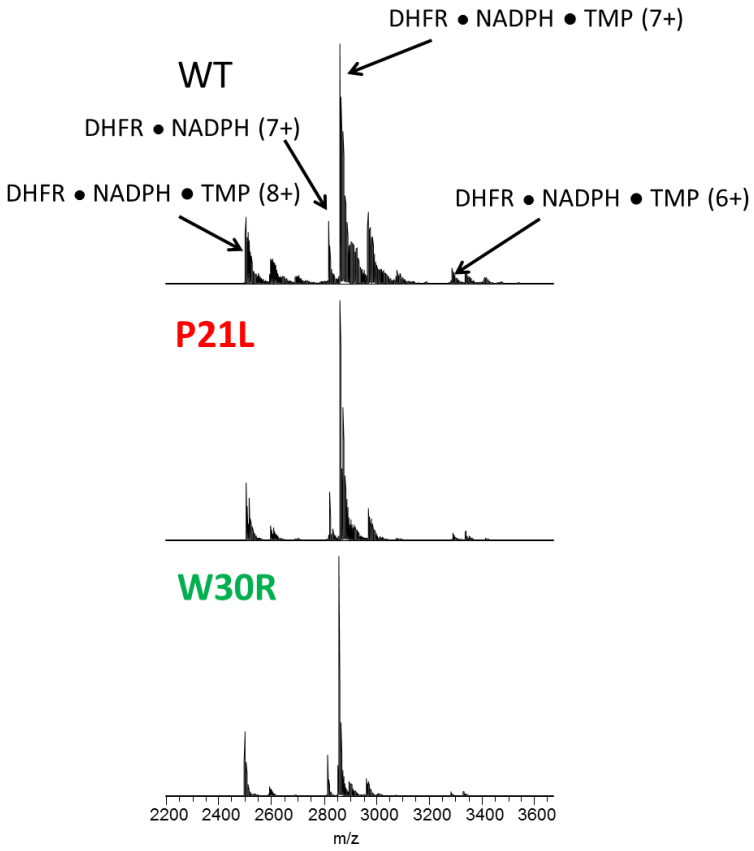

Figure 1. Native spray mass spectra for solutions containing each of three constructs of DHFR with TMP (2X) and NADPH (5X) in 150 mM ammonium acetate and 2% DMSO.

To shed light upon how the P21L and W30R single-point mutations cause TMP resistance, size exclusion chromatography coupled to native spray mass spectrometry was employed to monitor dissociation of the DHFR complexes as a means to evaluate relative k_{off} values (where k_{off} refers to a kinetic value that describes the rate of dissociation of the protein ligand complex).[35,36] Solutions containing various ratios and combinations of one DHFR variant (WT or P21L or W30R) with NADPH, DHR, MTX, and TMP were incubated, injected onto the SEC column, and the products were monitored. **Figure 2**

displays extracted ion chromatograms (XIC) for the resulting products for one representative solution containing DHFR, NADPH, and TMP.

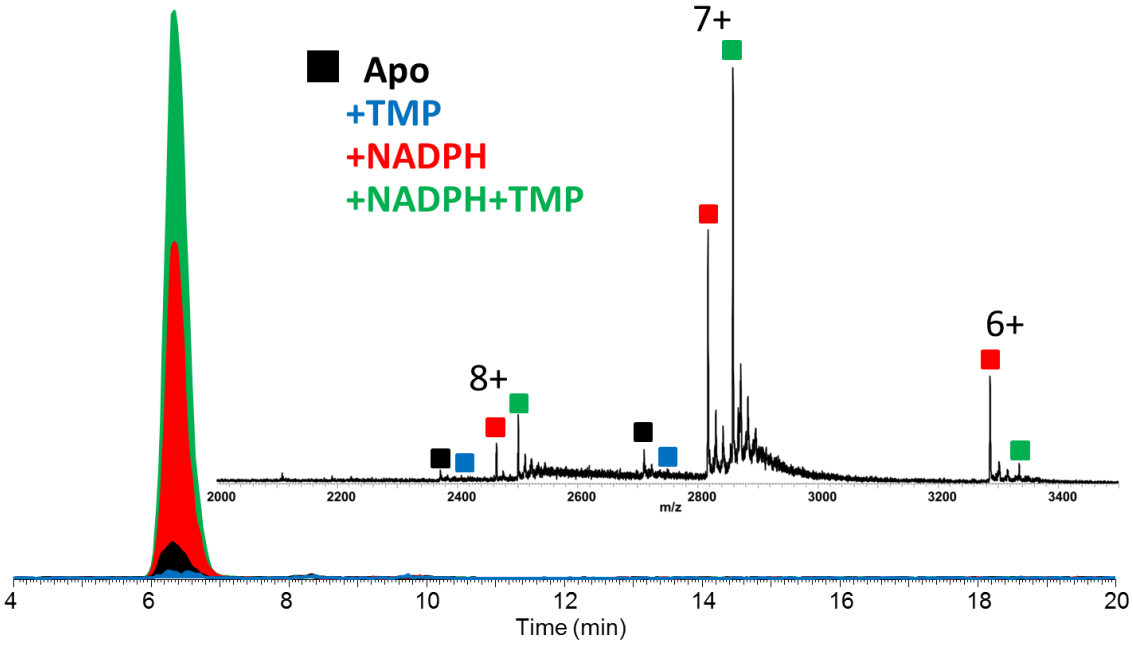

Figure 2. XICs of the protein complexes observed upon size-exclusion chromatography nativespray MS of a solution containing WT-DHFR, NADPH, and TMP (1:5:2 molar ratio with a protein concentration of 12 μM). Approximately 5 ug of protein was injected. The inset shows the averaged mass spectrum across the entire SEC peak.

Binary complexes (DHFR + TMP) and (DHFR + NADPH), and ternary complexes (DHFR + NADPH + TMP), as well as the ligand-free protein (apo DHFR) are observed, all in the low 6+, 7+, and 8+ charge states that are characteristic of nativespray conditions. The proportions of each species were determined from the SEC peak areas. The results for each of seven different solutions containing one of the three DHFR constructs and either no ligands, 2X DHF, 5X NADPH, 2X MTX, 5X NADPH + 2X MTX, 2X TMP, or 5X NADPH + 2X TMP are summarized in the histograms in **Figure 3.** The differences in NADPH complexation based on analysis of the solutions containing

DHFR + NADPH (5X) are striking, with P21L showing weaker binding (i.e. less abundant complexes) than WT followed by the W30R mutant. A less dramatic downward trend is also seen for formation of ternary complexes of DHFR+NADPH+TMP when comparing WT to both P21L and W30R.

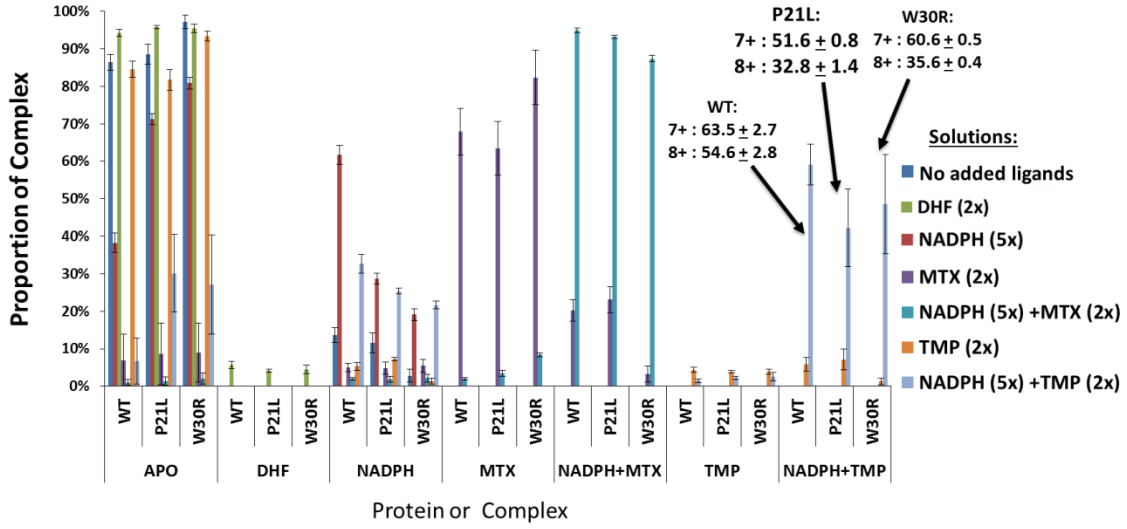

Figure 3. Distribution of complexes detected by native spray MS during SEC elution for each solution containing WT-DHFR, P21L, or W30R and various added ligands (NADPH, DHF, MTX, and TMP). The concentration of the protein in each solution was 12 μM.

In particular, the DHFR+NADPH+MTX complexes were lower in abundance for the W30R species relative to the WT and P21L variants. For solutions containing each of the three DHFR constructs and TMP, the abundances of binary (DHFR + TMP) complexes were very low, whereas the abundances of the corresponding ternary complexes (DHFR + NADPH + TMP) were ample. This interesting outcome suggests that the co-factor NADPH plays an important role in stabilization of the DHFR+TMP interactions.

The SEC-MS chromatograms were also used to monitor the profiles of unbound ligands, ones released from the protein complexes during their migration through the column (as well as those not bound in the initial incubate injected on column). Those

chromatographic peaks with long fronting tails are indicative of complexes with lower k_{off} values (i.e. slower dissociation of the protein-ligand complexes). SEC analysis of the incubates containing each protein construct and TMP with or without NADPH revealed that binary (protein + TMP) complexes were not detected for the W30R construct, thus indicating that the binary complexes, if formed, do not survive SEC separation. The binary (protein + TMP) complexes are detected for the WT and P21L constructs (**Figure 4a**). Interestingly, TMP is retained by W30R for ternary complexes containing NADPH (**Figure 4**), as also observed for the WT and P21L constructs.

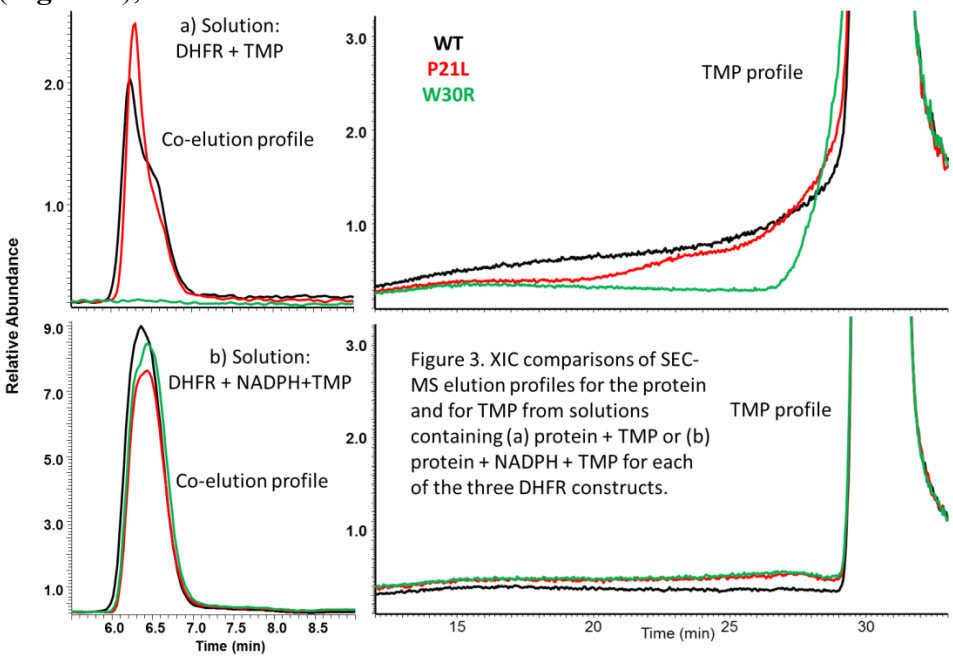

Figure 3. XIC comparisons of SEC-MS elution profiles for the protein and for TMP from solutions containing (a) protein + TMP or (b) protein + NADPH + TMP for each of the three DHFR constructs.

Figure 4. XIC comparisons of SEC-MS elution profiles for the protein and for TMP from solutions containing (a) protein + TMP or (b) protein + NADPH + TMP for each of the three DHFR constructs.

Monitoring the elution profiles of TMP shows distinctive front shoulders for the incubates containing WT or P21L and TMP, indicative of slow dissociation kinetics for the binary protein + TMP complexes. The fronting is virtually eliminated for solutions

containing each protein, NADPH and TMP, indicative of suppression of TMP release for all constructs. The same trend of differing k_{off} is witnessed for the substrate DHF as well (**Figure 5**). However, for DHF there is no protein associated peak, e.g. (DHFR + DHF), suggesting a faster k_{off} for DHFR+DHF complexes overall in comparison to the DHFR+TMP complexes.

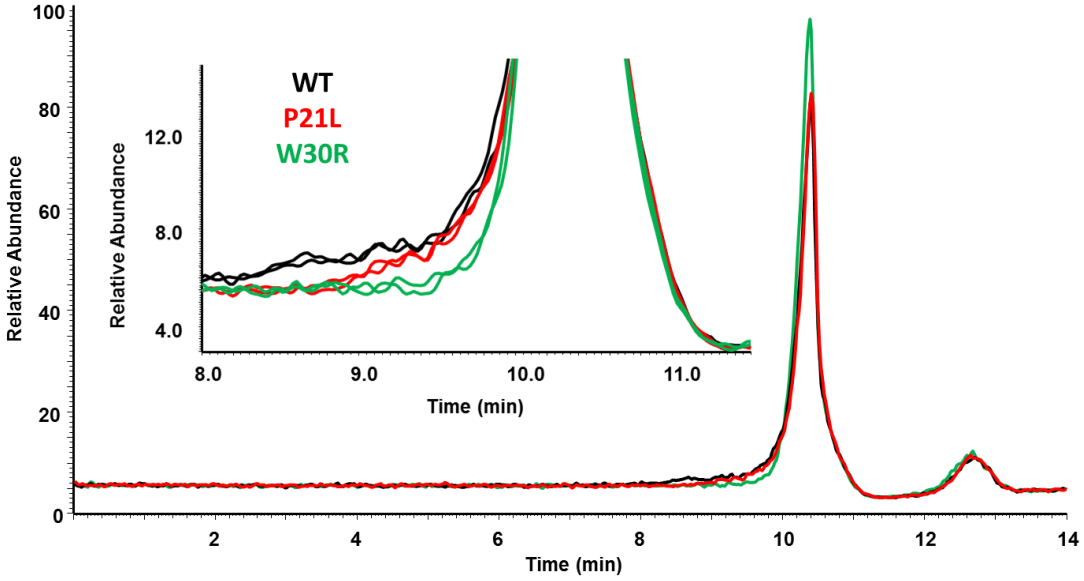

Figure 5. XIC comparisons of elution profiles for dihydrofolate (DHF) from solutions containing protein + DHF with respect to each DHFR construct (WT, P21L, W30R).

The negative control methotrexate was also evaluated in the same manner, and the results are highly complex (**Figure 6**) with multiple peaks in the protein eluting region (6 - 7 minutes) as well as later in the chromatogram. Due to this complexity, few conclusions can be extracted about these binary and ternary systems. Overall there is a significantly different elution time in the second bimodal chromatographic peak between the DHFR+MTX (**Figure 6a**) and the DHFR+NADPH+MTX (**Figure 6b**). This may be indicative of different dissociation pathways that are accessed between binary and ternary

complexes leading to the slower elution times for the ternary complexes seen in the later temporal profiles. Unfortunately, the novel propargyl-linked antifolates were too hydrophobic to elute from the column using native solvent conditions and thus could not be profiled in the same way by SEC-MS.

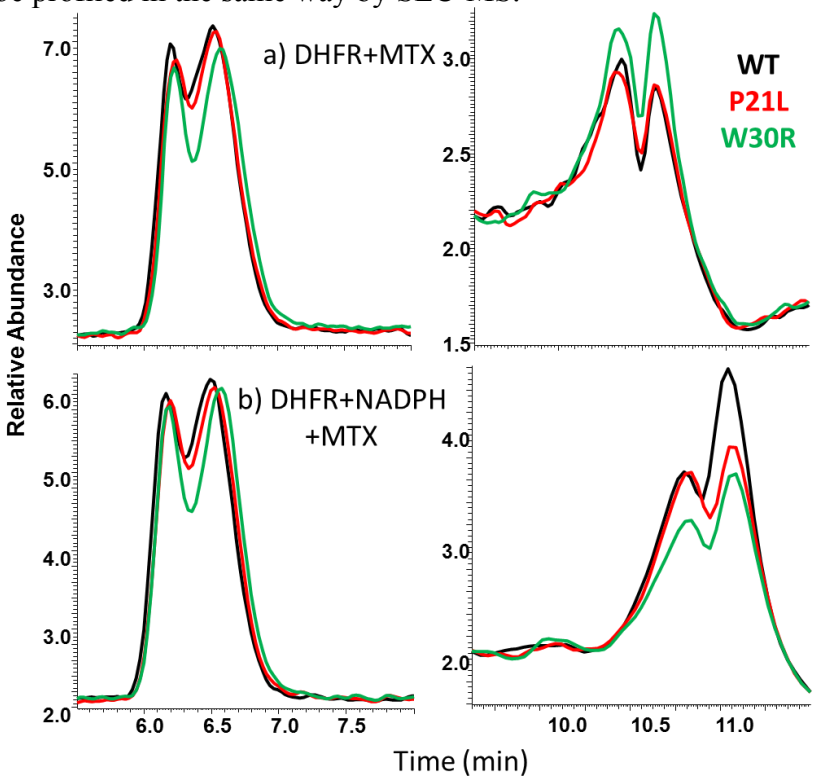

Figure 6. XIC comparisons of SEC-MS elution profiles for DHFR and MTX from solutions containing (a) DHFR+MTX or (b) DHFR + NADPH + MTX for each of the three DHFR constructs

Elution profiles for the binary (DHFR + NADPH) complexes and ternary (DHFR + NADPH + MTX) and (DHFR + NADPH + TMP) complexes are shown in **Figure 7 (a,b,c)**. Fronting occurs for each of the three protein constructs, and in a manner that is nearly identical which prevents detailed k_{off} comparisons. However, the release of the ligand can be monitored as shown in **Figure 7**. There are significant differences in the

release of NADPH in the presence or absence of the two inhibitors (MTX versus TMP) (**Figure 7a-c**) among the three protein constructs (**Figure 7d-f**).

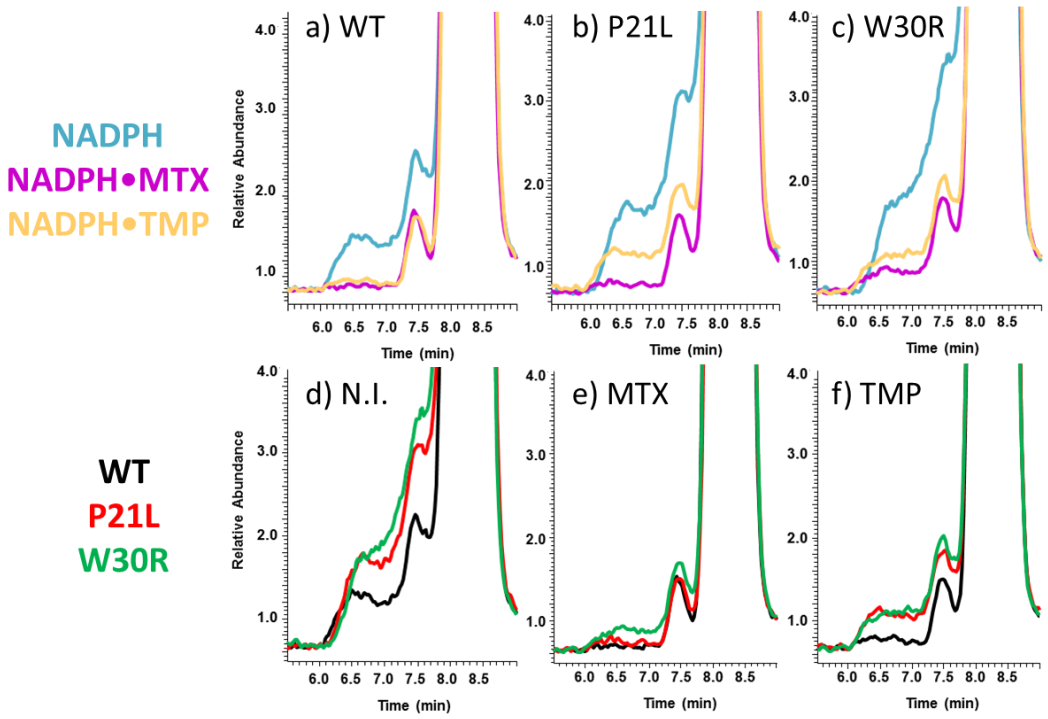

Figure 7. XIC traces (upper series) showing elution profiles of NADPH upon SEC-MS analysis of solutions containing protein+NADPH, protein+NADPH+MTX or protein+NADPH+TMP for each of the three DHFR proteins: (a)WT-DHFR, (b) P21L, and (c) W30R. XIC traces (lower series) showing elution profiles of NADPH upon SEC-MS analysis of solutions containing WT-DHFR, P21L or W30R with (d) no added inhibitor (N.I.), (e) addition of MTX, and (f) addition of TMP.

For WT DHFR, addition of either MTX or TMP significantly decreases the release of NADPH from the complexes (lower peak area in front of the bulk NADPH peak). This is interpreted to mean the complex formed has been stabilized thus deterring co-factor release during separation. The presence of MTX has a more notable impact on the release

of NADPH than does the presence of TMP for the P21L construct, and the release of NADPH is less substantially influenced by the presence of MTX or TMP for the W30R construct. These observations give evidence for how these single point mutations have allosteric effects on the co-factor binding and complex stabilization in addition to modulation of the binding of substrate and inhibitors.

The experiments described above show how native mass spectrometry coupled to SEC can be used to evaluate some of the factors that influence the relative stabilities of complexes and kinetic factors of complexation. With respect to examination of variations in conformation of the resulting complexes, UVPD can be used to decipher structural differences in the complexes based on variation in the extent of fragmentation along the backbone of each protein (with or without bound ligands). The UVPD fragmentation yields reflect the efficiencies of backbone cleavages along the protein. An example of a typical UVPD mass spectrum, the sequence coverage map constructed from the UVPD mass spectrum, and the relative extent of backbone cleavage at each position is shown for P21L DHFR protein in complex with NADPH in **Figure 8**. The UVPD mass spectra for each protein complex are used to create difference plots which allow facile comparison of the changes in backbone cleavage efficiencies between proteins or their complexes. An example is shown in **Figure 9a** for the three apo-proteins, in which the variations in UVPD fragmentation yields are shown from the N-terminus to the C-terminus for each of the P21L and W30R constructs relative to WT DHFR. The difference plots in **Figure 9a** diverge considerably for the two mutant constructs, and the most significant variations are color-coded on the protein structures in **Figure 9b,c**. Increases in the efficiency of backbone cleavages of the P21L or W30R complexes relative to WT DHFR upon UVPD are highlighted in red and decreases in the efficiency

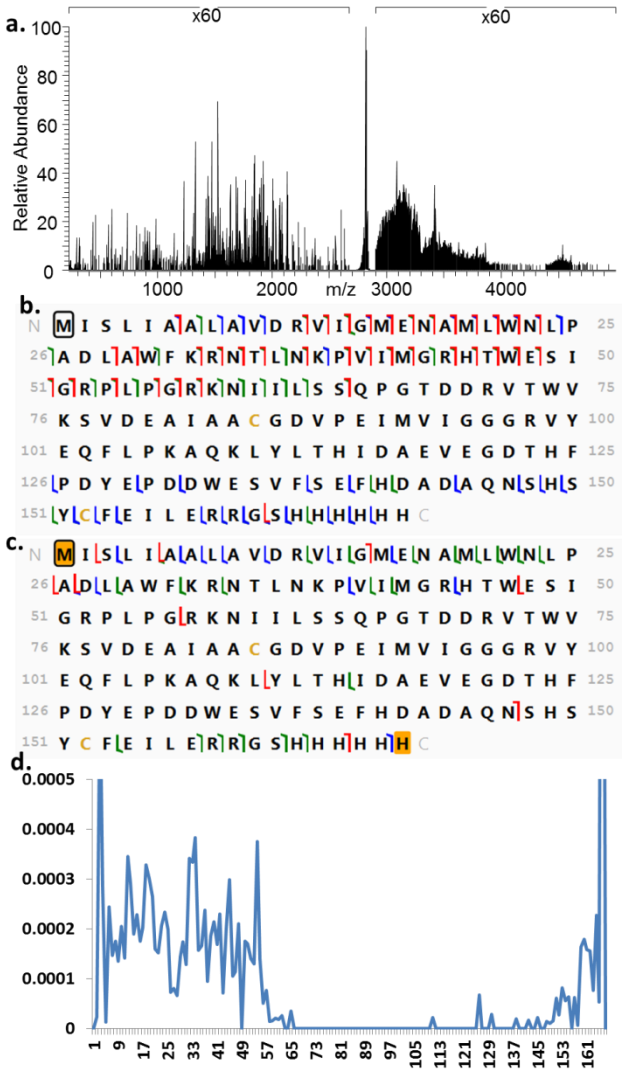

Figure 8. a) Fragmentation of the 7+ P21L DHFR+NADPH protein complex with the resulting b) apo ion identification and c) NADPH holo ion identificaiton (+745.0906). The sum of the intensities of both apo and holo ions from this complex can be seen in panel (d).

of backbone cleavage are highlighted in blue. Increases or suppression of backbone cleavages are presumed to indicate regions of the protein that become more flexible or more stabilized due to conformational changes, variations in intramolecular interactions, or other factors. There are relatively modest changes in UVPD fragmentation for the

P21L mutant relative to WT-DHFR; however the few changes that are observed are located proximal to the M20 loop or in the loop regions surrounding the substrate binding pocket.

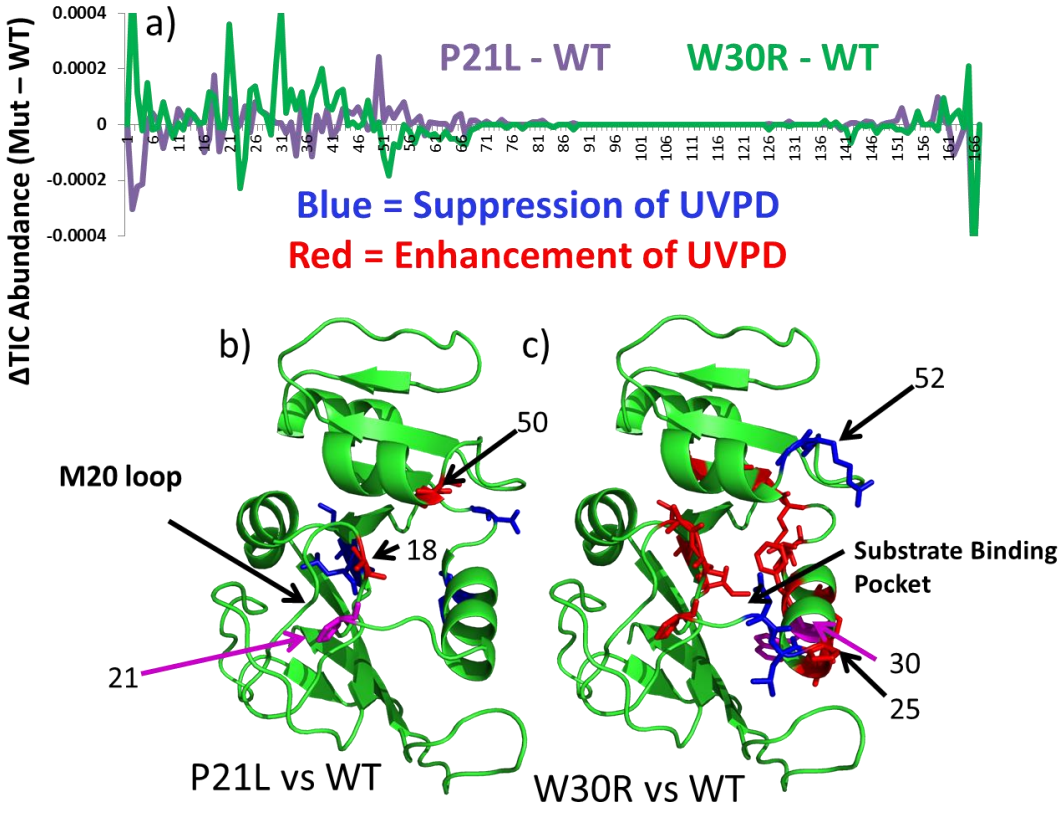

Figure 9. (a) Differences in UVPD fragmentation efficiencies for P21L (in purple) and W30R (in green) relative to WT-DHFR. Large differences in UVPD (those with >0.0001 ΔTIC abundance) are highlighted on the DHFR crystal structure (1RX3) for P21L (b) and W30R (c). Those residues highlighted in blue sticks show UVPD suppression, and those residues highlighted in red sticks show UVPD enhancement. Hot pink indicates the mutated residue.

More substantial variations in UVPD fragmentation occur for the W30R mutant, particularly shifts in backbone cleavage efficiencies in the regions of the substrate/inhibitor binding pocket as well as the M20 loop. Residue 21 of DHFR does not

play a major role in the core alpha helix/beta sheet composition of DHFR but rather occupies a peripheral position. This suggests that the P21L point mutation is less likely to cause a significant conformational re-organization of DHFR and instead the mechanism of TMP-resistance may originate from a kinetic or thermodynamic modulation in the uptake of TMP. In contrast, residue 30 is a key amino acid in the core structure of DHFR, and the W30R mutation replaces a hydrophobic amino acid with a more compact hydrophilic residue. The W30R mutation is anticipated to modulate TMP binding.

Following the use of UVPD-MS to examine the variations in UVPD fragmentation efficiencies of the ligand-free proteins, each of the binary complexes (+NADPH, +DHF, +MTX, +TMP, +1038, +1103) and ternary (NADPH + MTX, NADPH + TMP, NADPH + 1038, NADPH + 1103) complexes for each DHFR construct were examined by UVPD. The differences between each binary and ternary combination were calculated and summed for three key regions of the protein (M20 loop comprised of residues 10-24, substrate binding region consisting of residues 25-55, and C-termini region containing residues 126-155 which encompasses the G-H loop essential for protein activity) (plotted as heat maps in **Figure 10**). UVPD yielded low or no fragmentation of the backbone for one other essential loop[44] (F-G loop, spanning residues 115-130 in the middle of the protein) and thus could not be evaluated. Cumulative summation of the differences in UVPD of the binary or ternary complexes allows a convenient way to showcase the most significant changes in fragmentation for the large array of binary and ternary complexes and thus probe the complete conformational space of each mutant and ligand combination.

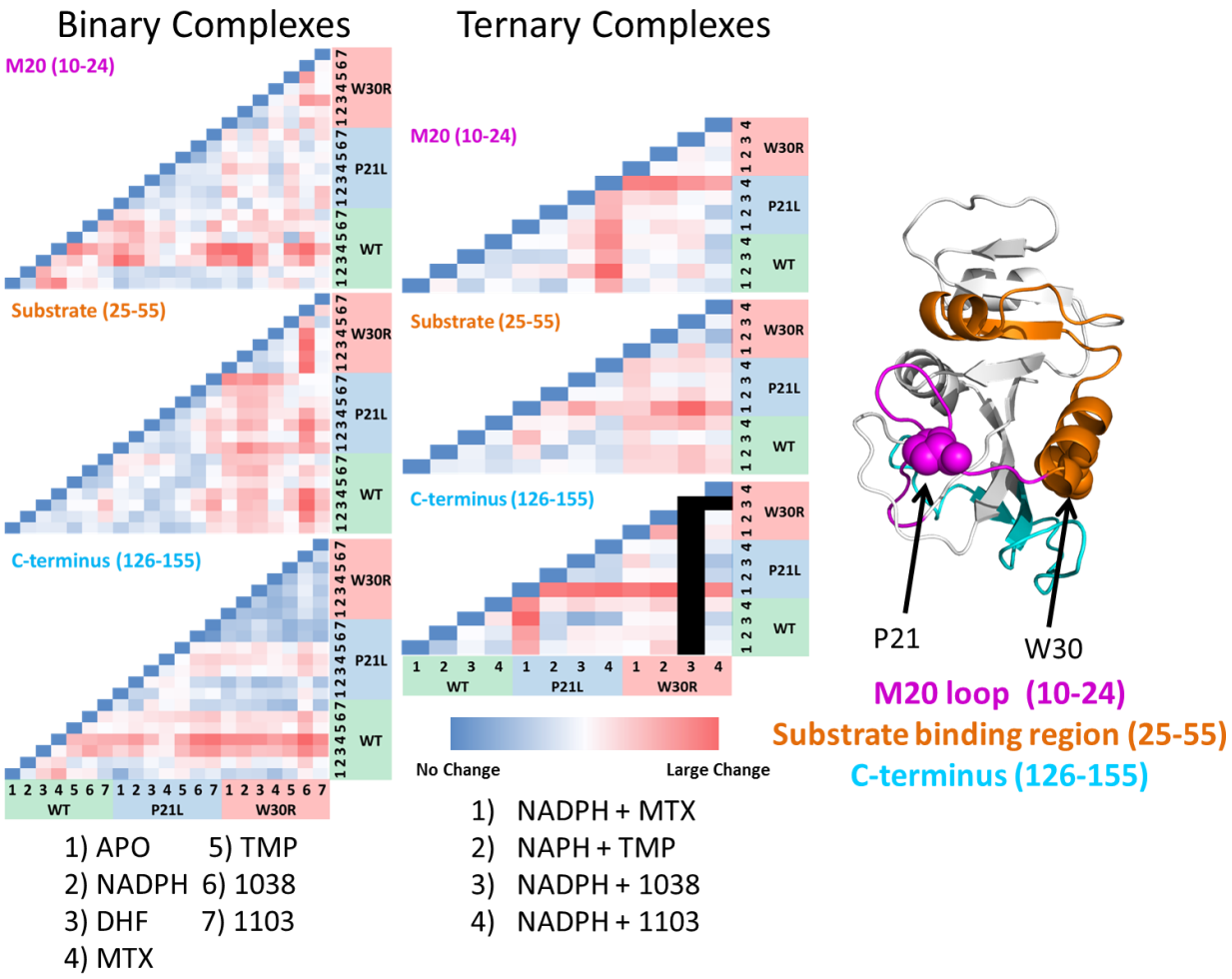

Figure 10. Direct comparison of absolute differences in UVPD fragmentation efficiencies between each combination of (a) binary and (b) ternary complexes for three regions of interest. These regions are the M20 loop (residues 10-24, hot pink), substrate binding region (residues 25-55, orange) and the C-terminus (residues 126-155, turquoise). These regions are highlighted on the DHFR crystal structure (1RX3) with specific residues P21 and W30 marked in spheres. Each number on the edge of the plot is related to the numbering system below it and separated into three different sections where green is WT, blue is P21L and red is W30R. Black bar in ternary complex of W30R with compound 1038 indicates that no ions in this region were detected.

For this phase of the study, two propargyl-linked antifolate inhibitors (1038 and 1103, structures shown in **Figure 11**) were also included to extend the strategy from well-characterized inhibitors to newly emerging candidates.

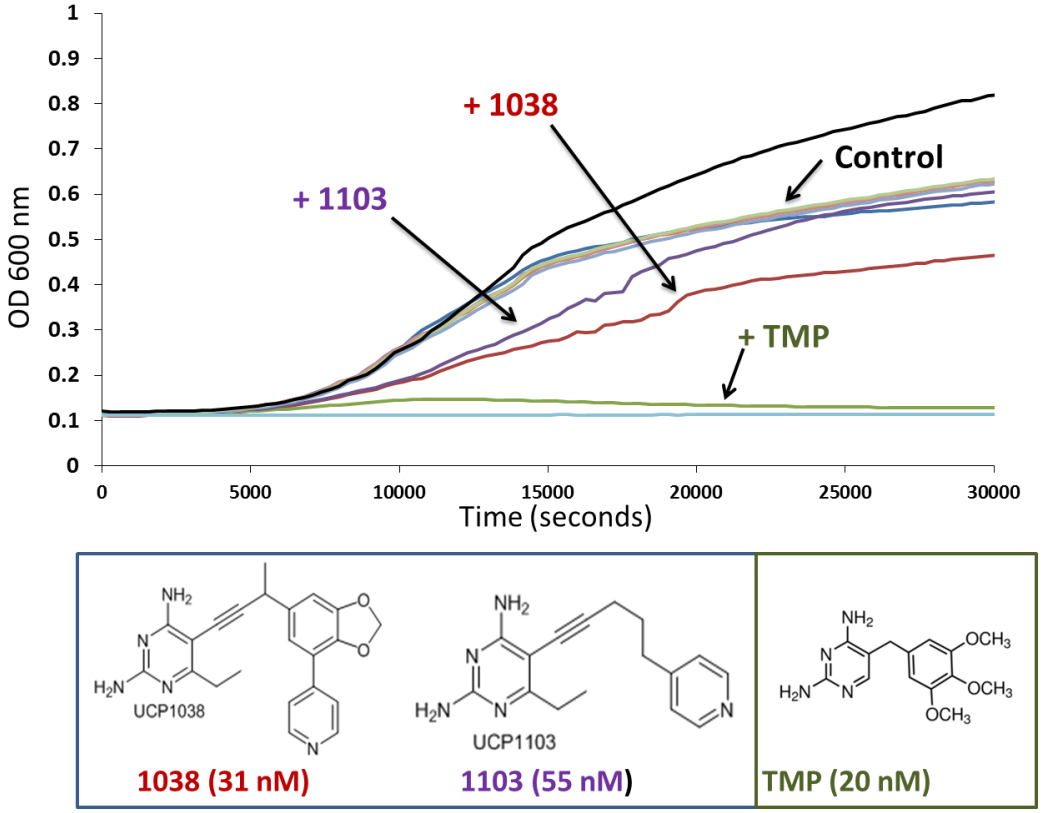

Figure 11. Growth curves of *E. coli* MG1655 cells incubated with 1 µg/ml inhibitor (compound 1038,1103, or TMP) relative to the control (no inhibitor added). The structures and IC_{50} (nM) values for four inhibitors of WT DHFR are shown. Multiple curves surrounding the control case represent other tested inhibitors that did not impede cell growth.

The largest consistent changes in UVPD fragmentation occurred over the substrate binding pocket region for W30R in both the binary and ternary complexes relative to all other ligand and mutant combinations. In particular a significant difference was observed for all W30R complexes containing DHF and MTX relative to the WT-DHFR complexes. Additional variations, although not as substantial, were found for the M20

loop and C-terminal region of the P21L binary complexes. Interestingly, these two regions of the protein interact with each other (particularly the M20 and G-H loops) during the enzymatic function of the protein, so it is encouraging that these trends are reflected in the UVPD trends through both regions of the protein[44]. For the ternary complexes, the P21L mutant exhibits the most significant variations in UVPD fragmentation in the M20 loop and C-terminal regions. The UVPD trends for the ternary complex (P21L + NADPH + 1103) suggests a different conformational change than observed for the corresponding WT and W30R complexes, and the ternary complex (P21L + NADPH + MTX) also displays a significant change in UVPD fragmentation efficiency in the C-terminal region. It is especially interesting to note that based on the UVPD trends, the P21L construct, unlike the W30R construct, does not seem to perturb the substrate binding pocket and instead specifically affects the activating loops of the protein.

To support the results derived from the SEC-MS and UVPD-MS methods, classical Michaelis-Menten kinetic experiments were performed (**Table 1**) as well as inhibitory kinetic testing (**Table 2**). The activity of each DHFR variant was evaluated via UV-Vis measurements based on monitoring the conversion of NADPH to $NADP^+$ in the presence of substrate DHF by UV-Vis at 340nm (**Table 1**).

Enzyme	Km [DHF] (μM)	Km [NADPH] (μM)	Kcat (s^{-1})	Kcat/Km [DHF] ($\mu M^{-1} s^{-1}$)	Kcat/Km [NADPH] ($\mu M^{-1} s^{-1}$)
WT	3.49 ± 0.53	4.95 ± 0.66	27.2 ± 2.5	7.79 ± 1.38	5.49 ± 0.89
P21L	1.86 ± 0.21	7.13 ± 1.76	28.4 ± 0.9	15.3 ± 1.79	3.98 ± 0.99
W30R	1.18 ± 0.32	0.60 ± 0.17	2.51 ± 0.04	2.12 ± 0.57	4.19 ± 1.2

Table 1. Michaelis-Menten kinetic parameters for the activity of each protein in the presence of substrate DHF.

	WT		W30R		P21L	
Compound	IC_{50} (nM)	K_i (nM)	IC_{50} (nM)	K_i (nM)	IC_{50} (nM)	K_i (nM)
TMP	20.4 ± 2.3	0.69 ± 0.13	478 ± 26	5.57 ± 0.69	220 ± 19.6	4.02 ± 1.14
MTX	17.5 ± 0.9	0.59 ± 0.09	11.2 ± 0.3	0.13 ± 0.02	15.1 ± 0.5	0.28 ± 0.08
UCP1038	31.0 ± 7.1	1.04 ± 0.28	130 ± 11.5	1.52 ± 0.21	38.2 ± 3.1	0.70 ± 0.19
UCP1103	54.7 ± 1.4	1.84 ± 0.28	296 ± 3.45	3.45 ± 0.39	117 ± 9.2	2.14 ± 0.60

Table 2. Inhibitory kinetic parameters

Rates of decreasing absorption at 340nm at varying concentrations of DHF and NADPH were used to calculate Michaelis-Menten parameters for K_m values for DHF and NADPH, respectively. Both variants (W30R and P21L) retained activity for conversion of DHF to THF. As summarized in **Table 1**, both the P21L and W30R constructs displayed relatively stronger interactions with DHF relative to WT DHFR based on their larger K_m values. While P21L displayed significantly weaker interactions for NADPH based on the K_m and W30R showed an order of magnitude stronger interaction NADPH than the WT. However, this factor does not seem to perturb the catalytic constant (K_{cat}) of the P21L protein which displays a turnover rate similar to that of the WT protein. The W30R construct exhibits an order of magnitude decrease in activity (K_{cat}) relative to WT DHFR in part due to the stronger interaction of the co-factor (NADPH) and its products. Upon comparing the efficiency of the enzyme activity (kcat/km) for both substrate and co-factor it was seen that the P21L construct is nearly twice as efficient and the W30R was a third as efficient in comparison to the WT, with respect to DHF. While measuring the efficiency with respect to NADPH the WT is the most efficient while P21L and W30R remain comparable with slightly lower efficiency. Overall, this suggests that the DHF binding and release for the full enzymatic process is most effected by the TMP resistant mutations.

While direct comparisons between K_m and the qualitative K_{off} values obtained from the SEC-MS measurements cannot be made, they provide complementary kinetic and thermodynamic information, respectively. K_m is indicative of protein-ligand binding (DHF and NADPH), enzymatic activity, and release of products (THF and NADP+) over the course of the reaction. Relative k_{off} displays information particularly about the rate of release of the unprocessed ligands (DHF and NADPH) or protein complex stability under study from a thermodynamic perspective and is related to the first binding steps in the enzymatic process. These are two independent parameters to consider when describing the enzymatic landscape. For instance, it was seen in general for the SEC-native MS case that protein-ligand complexes were most stable in the WT followed by the P21L and W30R constructs. Particularly for DHF dissociation, and binary TMP complexe dissociation (DHFR+TMP) the W30R was most unstable. However, inspecting the ternary TMP complex (DHFR+TMP+NADPH) the apparent k_{off} was nearly identical for the TMP elution(Figure 3b) as well as the NADPH elution (Figure 4f) upon comparison of P21L and W30R constructs. While P21L and W30R are more susceptibility for dissociation than the WT, these constructs display stronger affinity (lower Km values) with respect to DHF which may be reflective of lower energy dissociation pathway for substrate binding and subsequent product release. However, in the case of W30R the Kcat is diminished by an order of magnitude suggesting a significantly slower product release. Specifically by the tighter Km for the mutated W30R with respect to NADPH and the release of the NADP+ product. It is also seen that the P21L construct has the least affinity (Km) for NADPH, which is consistent with the location of mutation on the M20 loop which is responsible for NADPH binding.

Trimethoprim (TMP) is a well-known inhibitor of DHFR with a nanomolar binding constant (**Table 2**). IC50 values calculated for the two DHFR variants, W30R

and P21L, exhibited an order of magnitude increase in the inhibitory concentration compared to WT DHFR, thus confirming significant resistance to TMP. The complementarity between the relative thermodynamic k_{off} parameters is best revealed by comparing the IC50 and K_i values of the positive control inhibitor TMP. Additionally, the negative control (MTX) did not cause a large shift in neither the IC50 nor the K_i values for either of the two DHFR variants. The expectation for the P21L and W30R variants to release TMP more readily than would the WT construct based on its IC50 and K_i value, agrees well with the k_{off} SEC data (Figure 3). In addition, the fact that the K_i value is greater for W30R and P21L in comparison to the WT indicates that the mutant constructs are more susceptible to exchanging the inhibitor for the substrate. In short, the mutant constructs exhibit much less specificity overall for TMP than the WT proteins. The outcomes from the inhibitory kinetics of TMP align well with the k_{off} information determined from the earlier SEC-MS measurements, as well as with the structural insight from the UVPD trends. These trends indicated that there was a much broader conformational change in the binding pocket of W30R in comparison to the WT and P21L proteins. The conformational changes seen in the substrate binding pocket may facilitate this greater exchange for the substrate and the doubled IC50 for TMP in comparison to the P21L construct. Large conformational changes were not seen in the substrate binding pocket for P21L in comparison to the WT. P21L showed significant conformational change witnessed in the ternary complexes, specifically DHFR+NAPDH+MTX and DHFR+NADPH+1103 in both the M20 loop and c-terminal portion of the protein. The above evidence (amongst the other evidence) is suggestive that the two mutations induce resistance through two separate mechanisms.

To further parse out differences in mechanisms of resistance (and inhibition) as well as discover inhibitors for these TMP resistant mutants several novel propargyl-

linked antifolates (PLAs) were investigated. A small panel (6) of novel candidates was tested against wild type *E. coli* MG1655 cells to determine which ones exhibit positive inhibitory activities in vivo (**Figure 11**). Two of the candidates, 1038 and 1103, caused a significant decrease in cell growth at 1 μg/ml. The TMP control quenched nearly all cell growth at 1ug/ml, whereas the negative control MTX showed no inhibition of cell growth. It has been reported at that MG1655 cells have efflux pumps which render MTX ineffective *in vivo*.[45] Upon inspection of **Table 2**, compound 1103 resulted in large increases in the IC50 values for both W30R and P21L mutants relative to the WT protein. Compound 1038 was a reasonably effective inhibitor of P21L relative to the WT protein with small increase (more potent) in the K_i value and similar IC50s. The W30R mutant remained resistant to the new candidate inhibitors, possibly owing to the removal of a large hydrophobic residue (tryptophan) from DHFR in the substrate binding pocket which would be expected to interact with the hydrophobic propargyl linker portion of the inhibitor. These results further convey the large differences in the modes of action caused by the two point mutations of DHFR.

5.5 CONCLUSION

Using two separate mass spectrometry centric experiments the structure-function relationship of TMP resistant mutants were successfully evaluated revealing two different mechanisms of resistance one which directly modulate the core structure of the protein (W30R) while the other instills resistance through a more subtle manner adjusting the rigidity of the M20 loop (P21L) to aid in TMP release. These mass spectrometry results were supported by complimentary Michealis-Menten kinetic and respective inhibitory constants (IC50 and K_i) of each evaluated inhibitor. It was found that overall the qualitative k_{off} SEC experiments agree the best with the IC50 and K_i values. Perhaps even

more exciting is that compound 1038 which has shown to be previously very potent for *E. Coli* DHFR is also as potent for the P21L TMP resistant construct. These results show continued promise for propargyl-linked antifolates as effective inhibitors against pathogenic DHFR protein species.

5.6 REFERENCES

(1) Cullen, I. M.; Manecksha, R. P.; McCullagh, E.; Ahmad, S.; O'Kelly, F.; Flynn, R.; McDermott, T. E. D.; Murphy, P.; Grainger, R.; Fennell, J. P.; Thornhill, J. A. *Ir. J. Med. Sci.* **2012**, *182* (1), 81–89.

(2) Duffy, M. A.; Hernandez-Santiago, V.; Orange, G.; Davey, P. G.; Guthrie, B. *Br J Gen Pr.* **2013**, *63* (609), e238–e243.

(3) Pallett, A.; Hand, K. *J. Antimicrob. Chemother.* **2010**, *65* (suppl 3), iii25–iii33.

(4) Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119–140.

(5) Huovinen, P. *Antimicrob. Agents Chemother.* **1987**, *31* (10), 1451–1456.

(6) Toprak, E.; Veres, A.; Michel, J.-B.; Chait, R.; Hartl, D. L.; Kishony, R. *Nat. Genet.* **2012**, *44* (1), 101–105.

(7) Watson, M.; Liu, J.-W.; Ollis, D. *FEBS J.* **2007**, *274* (10), 2661–2671.

(8) Bryce, A.; Hay, A. D.; Lane, I. F.; Thornton, H. V.; Wootton, M.; Costelloe, C. *BMJ* **2016**, *352*, i939.

(9) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.

(10) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.

(11) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682.

(12) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.

(13) Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. *Nat. Commun.* **2013**, *4*, 1985.

(14) Zhou, M.; Wysocki, V. H. *Acc. Chem. Res.* **2014**, *47* (4), 1010–1018.

(15) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.

(16) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128–9135.

(17) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. *Nature* **2014**, *510* (7503), 172–175.

(18) Snijder, J.; van de Waterbeemd, M.; Damoc, E.; Denisov, E.; Grinfeld, D.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A.; Heck, A. J. R. *J. Am. Chem. Soc.* **2014**, *136* (20), 7295–7299.

(19) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3* (2), 126–132.

(20) Cong, X.; Liu, Y.; Liu, W.; Liang, X.; Russell, D. H.; Laganowsky, A. *J. Am. Chem. Soc.* **2016**, *138* (13), 4346–4349.

(21) Konermann, L.; Pan, J.; Liu, Y.-H. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234.

(22) Ben-Nissan, G.; Sharon, M. *Chem. Soc. Rev.* **2011**, *40* (7), 3627–3637.
(23) Holding, A. N. *Methods* **2015**, *89*, 54–63.
(24) Stefanowicz, P.; Kijewska, M.; Szewczuk, Z. *Anal. Chem.* **2014**, *86* (15), 7247–7251.
(25) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(26) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(27) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(28) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(29) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(30) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(31) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(32) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804–2812.
(33) Morrison, L. J.; Brodbelt, J. S. *Analyst* **2015**, *141* (1), 166–176.
(34) A. Loo, J.; A. Benchaar, S.; Zhang, J. *Mass Spectrom.* **2013**, *2* (Spec Iss).
(35) Bao, J.; Krylova, S. M.; Cherney, L. T.; Blanc, J. C. Y. L.; Pribil, P.; Johnson, P. E.; Wilson, D. J.; Krylov, S. N. *Analyst* **2015**, *140* (4), 990–994.
(36) Cherney, L. T.; Krylov, S. N. *Anal. Chem.* **2016**, *88* (7), 4063–4070.
(37) Muneeruddin, K.; Thomas, J. J.; Salinas, P. A.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (21), 10692–10699.
(38) Wang, G.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (15), 7293–7298.
(39) Flensburg, J.; Sköld, O. *Eur. J. Biochem.* **1987**, *162* (3), 473–476.
(40) Lombardo, M. N.; G-Dayanandan, N.; Wright, D. L.; Anderson, A. C. *ACS Infect. Dis.* **2016**, *2* (2), 149–156.
(41) G-Dayanandan, N.; Paulsen, J. L.; Viswanathan, K.; Keshipeddy, S.; Lombardo, M. N.; Zhou, W.; Lamb, K. M.; Sochia, A. E.; Alverson, J. B.; Priestley, N. D.; Wright, D. L.; Anderson, A. C. *J. Med. Chem.* **2014**, *57* (6), 2643–2656.
(42) Keshipeddy, S.; Reeve, S. M.; Anderson, A. C.; Wright, D. L. *J. Am. Chem. Soc.* **2015**, *137* (28), 8983–8990.
(43) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(44) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(45) Kopytek, S. J.; Dyer, J. C. D.; Knapp, G. S.; Hu, J. C. *Antimicrob. Agents Chemother.* **2000**, *44* (11), 3210–3212.

Chapter 6: Impact of G12 Mutations on the Structure of K-Ras Probed by Ultraviolet Photodissociation Mass Spectrometry

6.1 Overview

Native mass spectrometry combined with top-down ultraviolet photodissociation (UVPD) was employed to investigate the structural changes occurring from a somatic single residue mutation (G12X) of the GTP-ase protein K-Ras. Specifically, the prevalent and cancer-related protein K-Ras and its single residue mutants G12X (where X = C, V or S) were the focus of this study. G12X mutations of K-Ras are a prevalent oncogenic mutation. Complexes between K-Ras or the G12X mutants and guanosine 5'-diphosphate (GDP) or GDPnP (a stable GTP analog) were transferred to the gas phase by nano-electrospray ionization and characterized using UVPD. Variations in the efficiencies of backbone cleavages were observed upon substitution of GDPnP for GDP as well as for the G12X mutants relative to WT K-Ras. An increase in the fragmentation efficiency in the segment of the protein containing the first 50 residues was observed for the K-Ras/GDPnP complexes relative to the K-Ras/GDP complexes, whereas a decrease in fragmentation efficiency occurred in the section containing the last 100 residues. These stretches correspond to the phosphate and guanine binding regions, respectively, and are indicative of a change in the binding motif upon replacement of the ligand (GDP versus GDPnP). Unique changes in UVPD were observed for each G12X mutant with the cysteine and serine mutations exhibiting similar UVPD changes whereas the valine mutation was significantly different. The results support that variations in UVPD fragmentation reflect changes in conformations or dynamics arising from single amino acid mutations that modulate function of K-Ras.

6.2 INTRODUCTION

Proteins containing single amino acid mutations (presumably arising from somatic mutations acquired in a given cell over time) are expressed in a large number of cancerous tissues.[1,2] These mutations can specifically modulate the function of the protein in question, a factor which may promote cancer growth.[2] For example, single point mutations leading to constitutively active members of the rat sarcoma (Ras) family of proteins have been implicated in cell cycle progression, proliferation, apoptosis and senescence, all important in cancer.[3] Within the Ras family of proteins, K-Ras is the most frequently mutated isoform, present in 22% of all tumors analyzed and 90% of pancreatic tumors.[4] Of the point mutations found in K-Ras, 80% are found at codon 12. At this amino acid position, a Gly in wild-type (WT) K-Ras, mutations that introduce larger side chains block the ability of GTPase activating protein (GAP) to stimulate K-Ras hydrolysis of GTP, thereby locking it into an active state. Trapping the K-Ras protein in an active state leads to abnormally high concentrations of GTP-bound K-Ras, which results in upregulation of downstream pathways and unregulated cell proliferation and tumor growth. The particular amino acid substitutions found at the G12 position (designated here by G12X) vary in frequency, with G12D being the most prevalent, followed by G12V, G12C and others. These substitutions also vary in frequency by cancer type.[4] For example, the G12D mutation is often found in colon, lung, pancreatic and skin cancer tissues, among others. The G12V mutation is also widespread but can show more prevalence than G12D in skin cancers. The G12C mutation is extremely prevalent in lung cancers, and other G12X mutations, such as G12S appear less frequently.[4] There is growing evidence that, in addition to differences in distribution by cancer type, the identity of the side chain substitution at G12X leads to different downstream functional effects (vide infra). Therefore, development of novel tools for

investigating how particular mutations affect the conformation and binding interactions of K-Ras in particular, and other disease-relevant proteins in general, would be a compelling advance.

In recent years, mass spectrometry has become an increasingly popular approach for addressing a variety of questions in the arena of structural biology.[5–7] Strategies using chemical probes, including hydrogen/deuterium exchange (HDX) methods, to evaluate solvent accessibilities and map protein interfaces and other covalent-labeling methods have played a prominent role in advancing the applications of tandem mass spectrometry for studying structures of proteins.[8–12] These chemical-based methods have been used in the context of both bottom-up (proteolytic digestion of probe-modified proteins into peptides) and top-down (analysis of intact proteins) modes. For the probe-based methods, typically the solvent accessible regions of the proteins or protein complexes are reflected by the abundances of modified peptides (for the bottom-up methods) or the abundances of modified fragment ions (for top-down methods).[8–12] Perhaps the most information-rich approach entails HDX, in which the rate of exchange of the amide hydrogens of the protein backbone is correlated with the surface exposure of each residue.[8,11,12]

While chemical probe methods have been applied with great success to many systems, the use of top-down mass spectrometry to examine native-like intact proteins and protein complexes has accelerated, especially with advances in MS/MS methods that are sensitive to protein structure. Native mass spectrometry involves the use of buffered spray solutions containing volatile salts, typically ammonium acetate, which allows gentle transfer of proteins into the gas phase in low charge states (compared to proteins ionized using conventional nanospray conditions).[13–17] Although there are unresolved questions about the specific mechanisms of native-spray, it is believed that the charged

proteins remain to a large extent folded similarly to the tertiary and quaternary structures adopted in solution.[18–22] The shapes of proteins ionized by native-spray have been evaluated successfully by ion mobility mass spectrometry, in which the measured collisional cross sections have been correlated with the sizes of native proteins in solution, thus providing evidence that the proteins remain native-like during their transfer to the gas phase.[23–26] Cross-sections of proteins obtained from ion mobility measurements have been shown to increase significantly with charge state, thus signaling their unfolding.[18] Native MS has gained additional momentum with the growing availability of high resolution/high mass accuracy mass spectrometers that have allowed greater implementation of MS/MS methods for analysis of intact proteins.[27] To date, native mass spectrometry has been applied to a large array of applications in structural biology including examination of stoichiometries of protein complexes, ligand binding, binding/dissociation constants, conformational changes of proteins, and dynamic unfolding.[13–17,25–33]

With respect to MS/MS characterization of native-like proteins and protein complexes, electron-based activation methods, such as electron transfer dissociation (ETD)[34] and electron capture dissociation (ECD),[31,35–38] provide significant diagnostic sequence information. The abundances of the resulting fragments have been correlated with the degree of flexibility of different regions of the proteins, thus reflecting crystallographic B-factors.[31,37,39] Interestingly it has also been shown that salt bridges in proteins may survive upon electron transfer reactions, resulting in electron transfer but without dissociation of the resulting fragments, and the proteins release the electrostatically-bound fragment ions upon further activation.[34] Another activation technique, surface induced dissociation (SID), has proven to be particularly impressive for determination of quaternary arrangements of native multi-protein complexes.[40,41] A

third method takes advantage of high energy excitation via absorption of 193 nm photons to give the most extensive backbone fragmentation of proteins by any activation method.[42,43] Ultraviolet photodissociation (UVPD) has shown unprecedented sequence coverage for unfolded proteins as well as their natively-ionized analogs.[30,44,32,33] Additionally, the patterns of backbone cleavages promoted by UVPD not only correlate with the average B-factors of residues in proteins (which can reflect thermodynamic motions and/or heterogeneity of the crystal lattice among other factors) but also reveal ligand binding sites based on suppression or enhancement of fragmentation and retention of the ligand by the fragment ions upon protein dissociation.[30,33] The UVPD patterns have been shown to be modulated for different gas-phase conformers as measured by ion mobility mass spectrometry.[45,46] Most recently UVPD was used to elucidate ligand binding sites and monitor the conformational changes of dihydrofolate reductase (DHFR) upon binding of co-factor NADPH and its inhibitor methotrexate.[30] At the conclusion of this previous study it was proposed that UVPD should be tested on a protein with single amino acid mutations to evaluate the sensitivity of UVPD to subtle changes in sequence that might (or might not) cause conformational changes of the protein as well as induce alterations of ligand binding modes,[30] a proposal that catalyzed the present study.

In the present study, native spray mass spectrometry and top-down UVPD is used to characterize protein-ligand combinations of WT and clinically-relevant G12X mutants of K-Ras including G12C, G12S, and G12V. In our hands, the most common G12D variant was not stable upon purification (data not shown). These K-Ras variants are characterized either bound to guanosine diphosphate (GDP) or to guanosine 5'-[β, γ-imido]triphosphate (GDPnP), a non-hydrolyzable analog of guanosine triphosphate (GTP). The structural changes upon GTP hydrolysis of WT K-Ras have been studied using conventional biophysical methods and provide a comparative benchmark for the

gas-phase UVPD approach used in the present study.[47] We compare variations in UVPD fragmentation of GDP and GDPnP complexes containing WT K-Ras or the G12C, G12S and G12V variants to decipher unique structural changes arising from these single point mutations.

6.3 EXPERIMENTAL

6.3.1 Protein Expression and Purification

Recombinant human wild-type K-Ras (isoform 2, residues 1-169 (18,954 Da) was heterologously expressed from previously described expression plasmids[48] that also encode an *N*-terminal hexahistidine affinity tag followed by a tobacco etch virus (TEV) protease cleavage site preceding the rest of the sequence. Expression and purification for the G12C, G12S or G12V variants was the same. Briefly, the expression vector was transformed into *Escherichia coli* BL21 (DE3) and subsequent cultures were grown in Terrific Broth containing kanamycin (50 mg/L) with shaking (250 RPM) at 37 °C until the optical density$_{600nm}$ reached 0.4 - 0.6. Expression was induced through addition of isopropyl-β-D-thiogalactoside (IPTG, 0.5 mM)) and incubation was continued overnight (approximately 18 h) at 18 °C. Cells were pelleted by centrifugation at 11,325 × *g* for 20 min at 4 °C. The pellet was harvested and either used immediately for purification, or flash frozen in liquid nitrogen and stored at – 80 °C for later use. A fresh or thawed pellet was resuspended in 100 mL Lysis Buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole at pH 8) supplemented with protease inhibitor cocktail (Roche complete EDTA free). β-Mercaptoethanol (2 mM final concentration) was added to the solution and cells were lysed by sonication. Cell debris was pelleted by centrifugation at 39,191 × *g* for 20 min at 4 °C and discarded. The remaining supernatant was incubated for 1 h with Qiagen Ni-NTA resin slurry (6 mL) pre-equilibrated in Lysis Buffer. The slurry was then

transferred to a column and washed with Lysis Buffer (50 mL) supplemented with imidazole (20 mM). Hexahistidine-tagged K-Ras was then eluted using Lysis Buffer (10 mL) supplemented with imidazole (200 mM). The hexahistidine affinity tag was removed by treatment with *N*-terminal His-tagged TEV protease (1 mg protease for every 25 mg of total protein) along with supplementation of GDP (1 mg GDP for every 20 mg of total protein) and dialyzing overnight against Dialysis Buffer (20 mM Tris, 300 mM NaCl, 5 mM imidazole, 1 mM dithiothreitol (DTT) and 0.5 mM EDTA at pH 8). Following dialysis, the protein solution was diluted five-fold with Low-Salt Buffer (20 mM Tris, 50 mM NaCl at pH 8) and supplemented with $MgCl_2$ (5 mM) and GDP (1 mM). Uncleaved hexahistidine-tagged K-Ras and *N*-terminal His-tagged TEV protease were removed from the final solution by incubation (1 h) with Ni-NTA resin slurry (3 mL), pre-equilibrated with Low-Salt Buffer, followed by centrifugation to pellet and discard the Ni-NTA resin.

Each of the WT, G12C, G12S or G12V K-Ras variants was further purified by ion exchange chromatography using a Q-Sepharose Fast Flow column (8 mL, GE Healthcare Life Sciences). The column was washed using wash buffer (20 mM Tris, 100 mM NaCl at pH 8). K-Ras protein was eluted using Elution Buffer (20 mM Tris, 250 mM NaCl at pH 8). Remaining high molecular weight impurities were removed using an Ultracel 50 kDa molecular weight cutoff (MWCO) filter (Amicon Ultra 4 Centrifugal Filter). The filter flow-through containing K-Ras was exchanged into Storage Buffer (10 mM Tris, 25 mM NaCl at pH 8) using an Ultracel 10 kDa MWCO filter (Amicon Ultra 4 Centrifugal Filter) and concentrated to a final concentration of 1 - 2 mg/mL. Typical yields for each of the purified K-Ras variants were 5 - 10 mg / 1 L culture.

6.3.2 K-Ras Nucleotide Exchange

K-Ras variants (30 μM) were incubated in Exchange Buffer (25 mM Tris, , 10 mM EDTA, 1.5 mM guanosine 5'-[β, γ-imido]triphosphate (GDPnP, Sigma Aldrich), 1 mM dithiothreitol (DTT) at pH 8) for 1 h. $MgCl_2$ was subsequently added to the incubation to a final concentration of 25 mM, and the mixture was incubated for an additional 30 min. The protein was then exchanged again into Exchange Buffer using an Ultracel 10 kDa MWCO centrifugal filter (Amicon Ultra) and the above procedure was repeated a second time. In a final step, the protein was exchanged into a similar buffer that lacked EDTA (25 mM Tris at pH 8). This procedure loads K-Ras with the nonhydrolyzable GTP mimic, GDPnP and results in approximately 95% GDPnP loading as determined by MS (see below).

6.3.3 Mass Spectrometry

Equimolar protein/ligand solutions were prepared at a concentration of 15 μM at pH 7.8 buffered with 50 mM ammonium acetate with 5 μM magnesium acetate. The solutions were infused via a gold-coated static tip operated between 1.5 and 1.7 kV at a capillary temperature of 200°C. The pH was adjusted with ammonium hydroxide. All experiments were performed on a Thermo Scientific Orbitrap Elite mass spectrometer (Bremen, Germany) equipped with a Coherent Excistar 193 nm excimer laser (Santa Cruz, CA) and modified to allow UVPD as described previously.[42] All mass spectra were collected at 240K resolution at *m/z* 400. 50 scans were averaged for the MS1 spectra. All UVPD experiments were performed using a single 3 mJ pulse (without focusing or collimation of the laser). The 8+ charge state of each protein or protein/ligand complex was selected for MS/MS analysis using an isolation width of 12.5 m/z. The automatic gain control was set to achieve a signal level of 2e5 with a fill time of 1 second for UVPD experiments. The Orbitrap mass analyzer was scanned from m/z 220 to 4000, and

750-1000 scans were averaged for each UVPD mass spectrum. Three to four replicates were collected for each protein/ligand complex.

6.3.4 DATA ANALYSIS

The UVPD mass spectra were decharged using Xtract with a S/N of 2 and a fit factor of 44% and remainder of 25%. The monoisotopic ions were then searched against the respective K-Ras amino acid sequence using a version of Prosight PC 3.0 modified to accommodate UVPD fragmentation. This involves searching for the following ions: (a ,a•, b, c, x, x•, y, y-1, z). The spectra were also searched by considering all fragment ions with a mass shift corresponding to each ligand mass both with and without the coordinating Mg^{2+}. The more abundant holo fragment ion species contained the ligand as well as the coordinated divalent magnesium for both the GTP and GDPnP complexes of each K-Ras variant. GDPnP was used as a stable surrogate of GTP that does not undergo hydrolysis. Specifically fragment ion searches undertaken for GDP•Mg^{2+} complexes included the mass shift 462.9781-464.9937 Da and those undertaken for the GDPnP•Mg^{2+} complexes included the mass shift 541.9604-543.9760 Da. For analysis of backbone cleavage yields upon UVPD, the abundances of the holo fragment ions were collectively summed with the abundances of the corresponding apo ion series. All identified ions were normalized to the total ion current of the spectrum to allow direct comparison across all spectra. Identified ions (both apo and holo) from N- and C-terminal ions were summed together as described in Cammarata *et al.*,[30] unless stated otherwise. In short, fragment ions arising from cleavage of the backbone positions between pairs of adjacent amino acids in the protein sequence were collectively summed. For example, all N-terminal product ions (a_n, b_n, and c_n ions) arising from backbone cleavages that occur C-terminal to a specific amino acid were summed with all the C-terminal product ions (x_{R-}

$_{n+1}$, y_{R-n+1}, and z_{R-n+1} ions) arising from cleavages that occur N-terminal to the same amino acid, where n is the residue number and R is the total number of amino acids in the protein. This value is calculated for each amino acid to convey the backbone cleavage efficiency adjacent to each particular residue. The determination of analytical significance of a change in backbone cleavage efficiency upon UVPD is described in **Figure 1**.

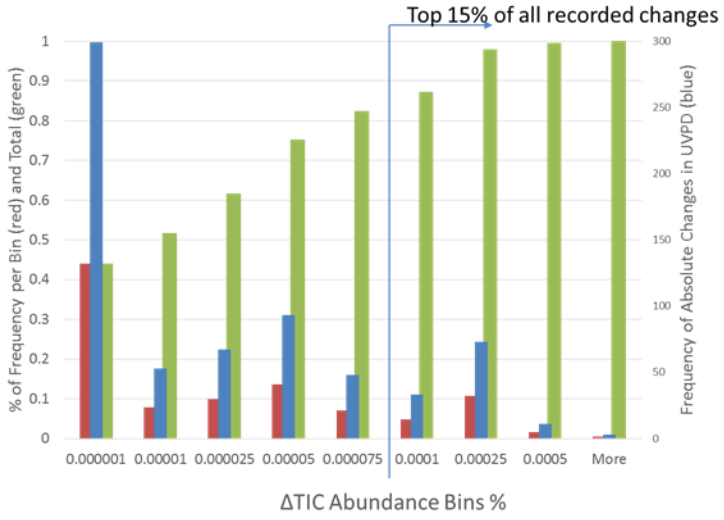

Figure 1. Histogram illustrating determination of significant changes in backbone cleavage efficiencies for all calculated changes in fragment ion abundances from the UVPD data. Blue bars show the frequency of bin level of the change in UVPD backbone cleavage efficiency (considering over 600 backbone cleavage efficiency values based on aggregation of all four proteins), and the red bars show the percent of the total population of backbone cleavages that correspond to the indicated changes. The green bars indicate the total additive percentage from small to large changes, thus illustrating that only the top 15% of changes in UVPD backbone cleavage efficiency are considered significant. Any change greater than an absolute signal change of 0.0001% (generally at least a factor of 2X or greater change in abundance) falls within the top 15 percent of all variations in backbone cleavages and is considered a significant change.

Crystal structures 4OBE and 3GFT were used for the WT GDP and WT GDPnP models, respectively. Diagrams of the crystal structure (3GFT) of WT K-Ras•GDPnP•Mg^{2+} are shown in **Figure 2** to illustrate the helices, loops, switches, and beta strands.

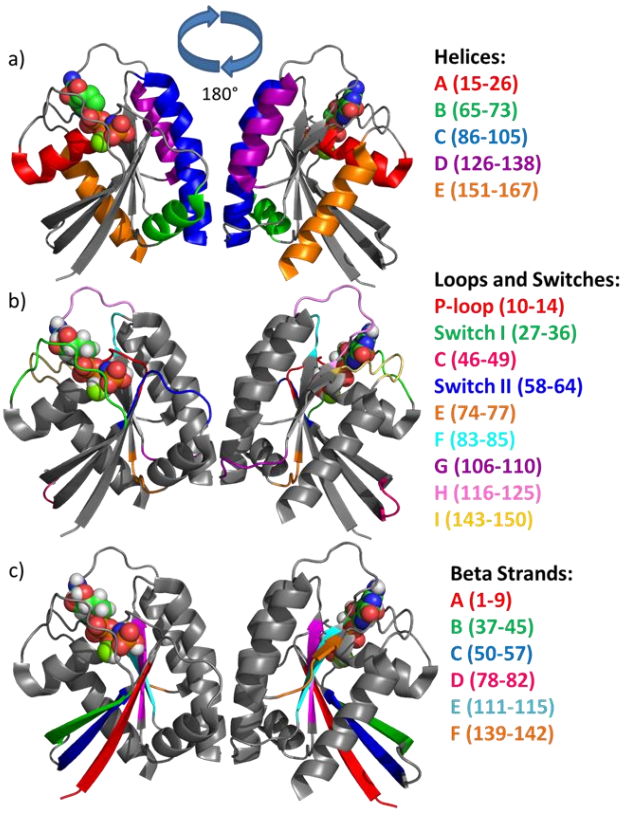

Figure 2. Color coded diagram for K-Ras showing various (a) helices, (b) loops and switches, and (c) beta strands based on the crystal structure 3GFT. The space-filled component represents GDPnP and Mg^{2+}. Front and back views are shown on the left and right. Note that only switch I, switch II, and loop P are formally named in the crystal structure. The other helices, loops, switches, and beta strands are labelled here to facilitate discussion of the structure.

6.4 RESULTS

Solutions containing each combination of K-Ras protein (WT, G12C, G12V, or G12S) and one ligand (GDP or GDPnP) were infused using native ESI conditions, and the 8+ charge state was subsequently isolated and subjected to 193 nm UVPD to produce informative fragmentation patterns. A representative ESI-mass spectrum is shown in **Figure 3a** for WT K-Ras, and the characteristic narrow and low charge state distribution (9+, 8+, 7+) is observed. The types of complexes formed upon addition of GDP or

GDPnP to the solutions are displayed in **Figure 3b-i,** showing the prominent formation of ternary complexes containing the protein, the nucleotide, and divalent magnesium.

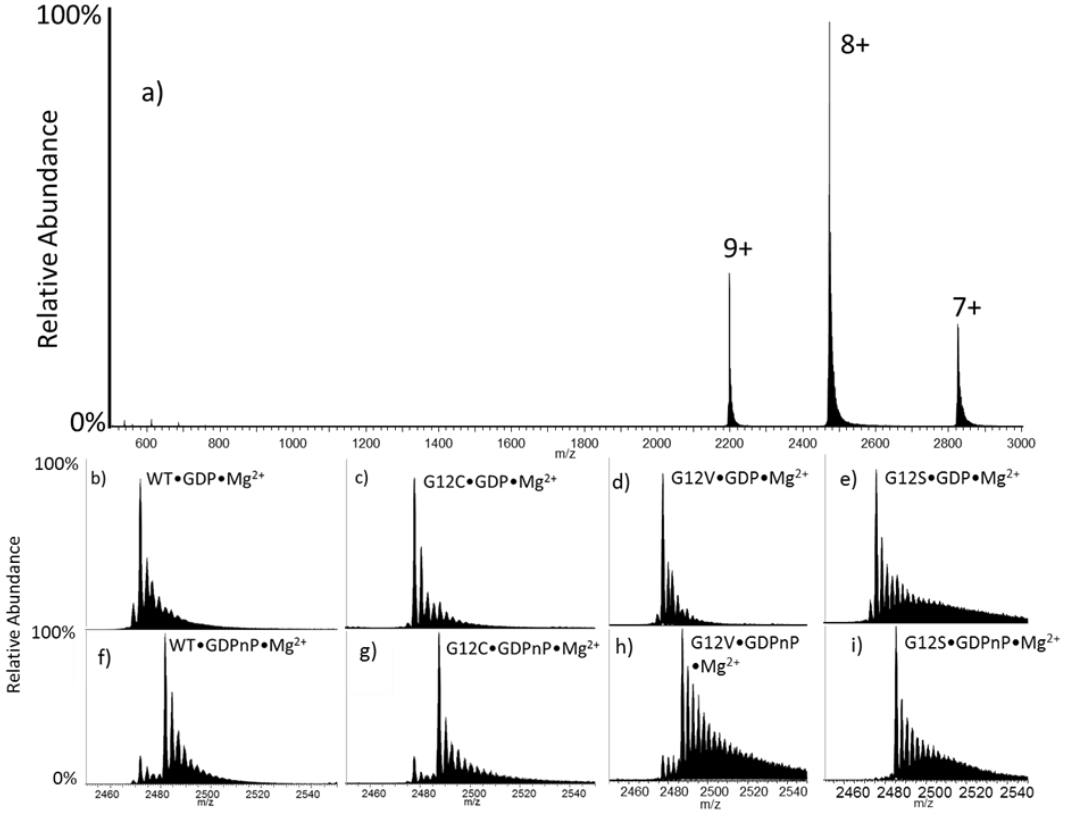

Figure 3. a) ESI mass spectrum of K-Ras (WT) sprayed from a 50 mM ammonium acetate solution (pH 7.8) and expanded regions covering the 8+ charge state for b) WT•GDP•Mg2+ , c) G12C•GDP•Mg2+ , d) G12V•GDP•Mg2+ , e) G12S•GDP•Mg2+ , f) WT•GDPnP•Mg2+ , g) G12C•GDPnP•Mg2+ , h) G12V•GDPnP•Mg2+ , and i) G12S•GDPnP•Mg2+.

The 8+ charge state was selected for UVPD owing to its large abundance. Replacement of GDP by GDPnP occurred with about 95% efficiency. Each complex of interest was isolated and activated by UVPD to generate the MS/MS spectra that are shown in **Figure 4**. The raw UVPD fragmentation patterns were decharged using XTRACT to produce the

deconvoluted spectra in **Figure 5**, with the masses of the fragments extending from very low mass up to 11,000 Da.

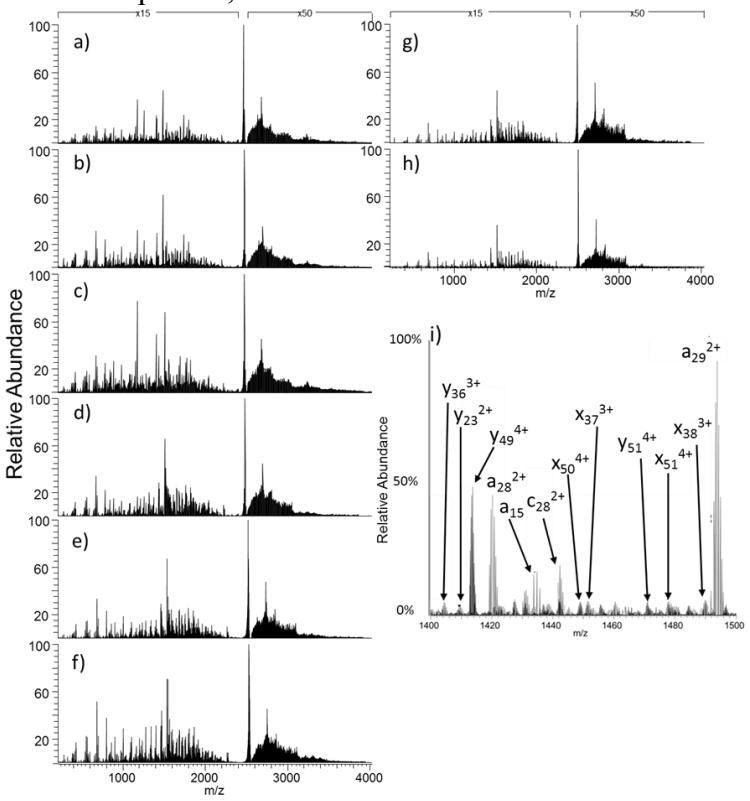

Figure 4. UVPD mass spectra of K-Ras (8+) activated using a single 3 mj pulse for a) WT•GDP•Mg2+, b) G12C•GDP•Mg2+, c) WT•GDPnP•Mg2+ and d) G12C•GDPnP•Mg2+ , e) G12V•GDP•Mg2+, f) G12V•GDP•Mg2+, g) G12S•GDPnP•Mg2+ and h) G12S•GDPnP•Mg2+. (i) There is also an expanded view from 1400-1500 m/z with selected and labelled ions from panel (a).

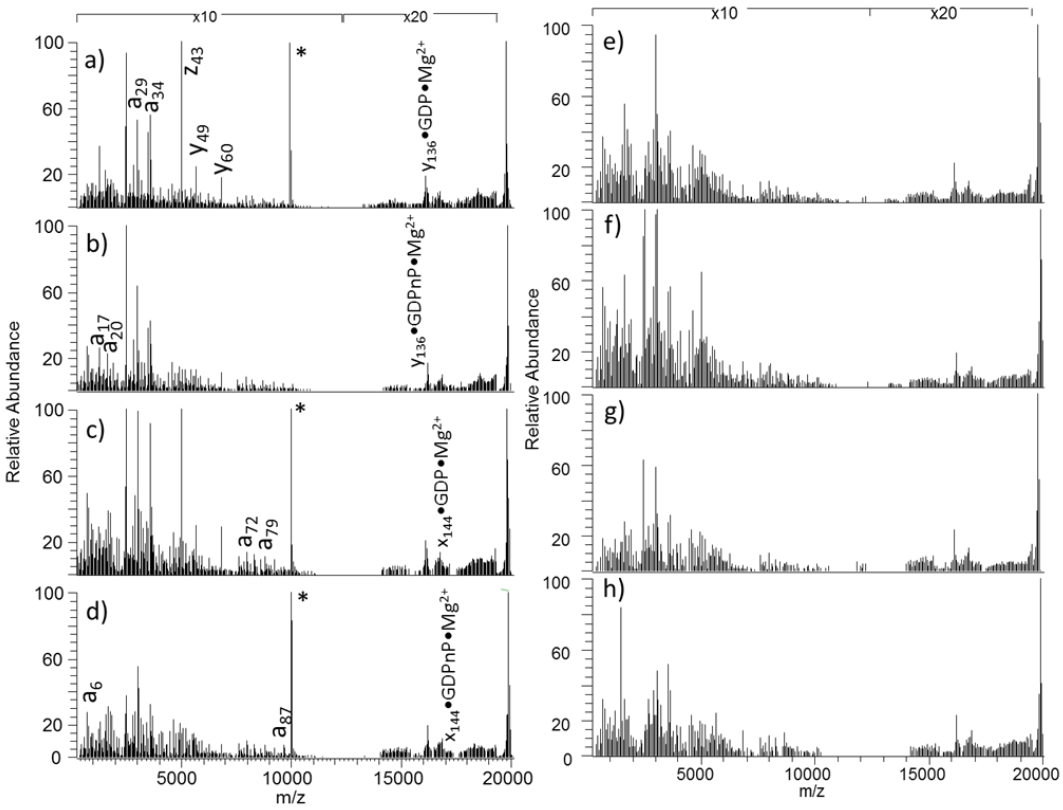

Figure 5. Deconvoluted UVPD mass spectra of WT K-Ras corresponding to the spectra in Supplemental Figure 4 for a) WT•GDP•Mg^{2+}, b) WT•GDPnP•Mg^{2+}, c) G12C•GDP•Mg^{2+} and d) G12C•GDPnP•Mg^{2+}, e) G12V•GDP•Mg^{2+}, f) G12V•GDPnP•Mg^{2+}, g) G12S•GDP•Mg^{2+} and h) G12S•GDPnP•Mg^{2+}. The asterisks indicate deconvolution artifacts at half the mass of the precursor.

Assignment of the fragment ions and interpretation of the data entailed analysis of both the apo (no ligand) and holo (retaining the ligand) product ions, as described in the experimental section. For this analysis, products containing just the nucleotide by itself or containing both the nucleotide and magnesium ion(s) were considered. Exhaustive searches revealed that more products were successfully identified containing both ligands (nucleotide plus Mg^{2+}) for all protein complexes of interest (**Figure 6** for WT), and thus these products were targeted in the present study. The product ions that were specifically searched for the protein/GDP complexes included those incorporating a mass shift of

462.9781-464.9937 Da (relative to the apo fragment ions), and those searched for the protein/GDPnP complexes included those with a mass shift of 541.9604-543.9760 Da. On average, approximately 142 GDP•Mg^{2+} holo ions and 139 GDPnP•Mg^{2+} holo ions were identified for each protein construct.

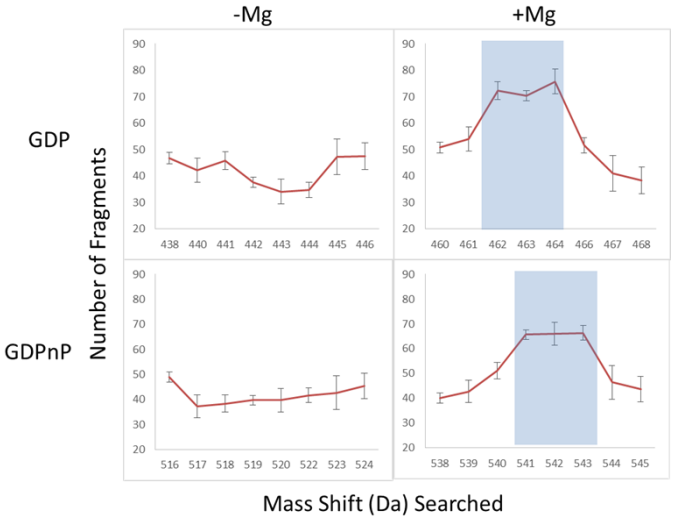

Figure 6. Deconvoluted UVPD mass spectra of WT K-Ras corresponding to the spectra in Supplemental Figure 4 for a) WT•GDP•Mg2+, b) WT•GDPnP•Mg2+, c) G12C•GDP•Mg2+ and d) G12C•GDPnP•Mg2+ , e) G12V•GDP•Mg2+, f) G12V•GDPnP•Mg2+, g) G12S•GDP•Mg2+ and h) G12S•GDPnP•Mg2+. The asterisks indicate deconvolution artifacts at half the mass of the precursor.

Overall, the net sequence coverage obtained by UVPD of the GDP complexes was 96% for each construct and 92% for the corresponding GDPnP complexes upon including both apo and holo fragment ions in the searches. Surprisingly, the sequence coverage obtained upon UVPD of the 20+ (denatured) charge state of WT K-Ras decreased to 55%. A similar drop in coverage has been noted previously upon comparison of UVPD mass spectra of native versus denatured proteins in this laboratory.[30,32]

6.5 DISCUSSION

6.5.1 UVPD of WT K-Ras Complexes: analysis of holo fragment ions and determination of ligand binding sites

Previously it was shown that analysis of the holo (ligand-containing) fragment ions produced upon UVPD of protein-ligand complexes allowed predictions about the ligand binding sites.[30,33]

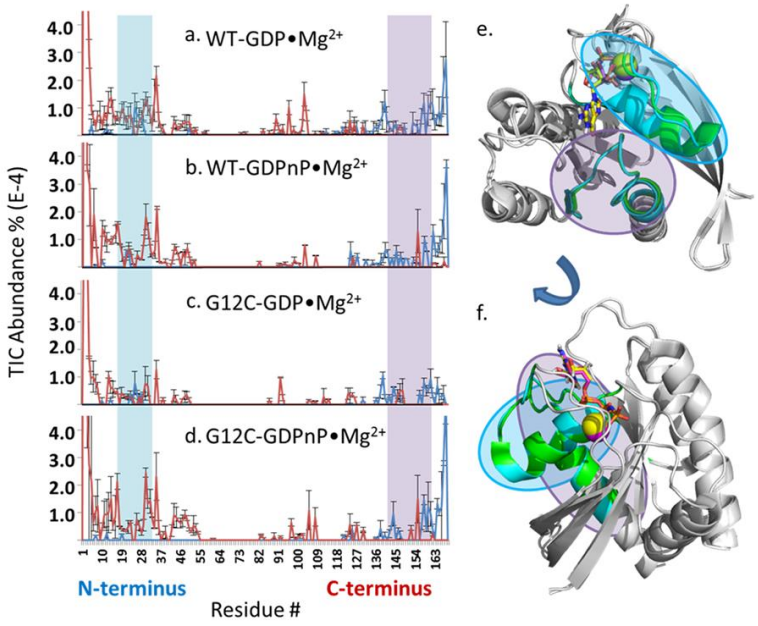

Figure 7. (a,b,c,d) Relative abundances of the holo (ligand-containing) fragment ions produced upon UVPD of WT K-Ras and G12C complexes, color coded as N-terminal holo fragments (blue traces) and C-terminal holo fragments (red traces). The UVPD fragment ion maps for the corresponding G12V and G12S complexes are shown in Supplemental Figure 7. The abundances are plotted relative to the amino acid sequence along the x-axis to convey the relative efficiencies of backbone cleavages adjacent to each residue. Shaded regions (blue and lavender shading) highlight the regions of the most notable changes in UVPD fragmentation efficiencies between the GDP and GDPnP complexes and are expanded in Supplemental Figure 8. (e and f) The same blue and lavender regions are highlighted on a superposition of crystal structures 3GFT (WT K-Ras•GDPnP •Mg^{2+} complex with green for WT K-Ras features and yellow for the GDPnP ligand) and 4OBE (WT K-Ras•GDP•Mg^{2+} complex with turquoise for WT K-Ras features and pink for the GDP ligand). Two rotated views are shown in e and f.

This analysis focused on identification of those stretches of the protein for which both N-terminal and C-terminal holo fragment ions overlapped, thus indicating regions of the protein that interacted with the ligand(s) and retained the ligand during fragmentation of the protein. Holo fragment ion plots for each protein/ligand combination of WT and G12C K-Ras are displayed in **Figure 7,** and the holo fragment ion plots for the corresponding G12V and G12S complexes are shown in **Figure 8** with expansions of selected regions in **Figure 9.**

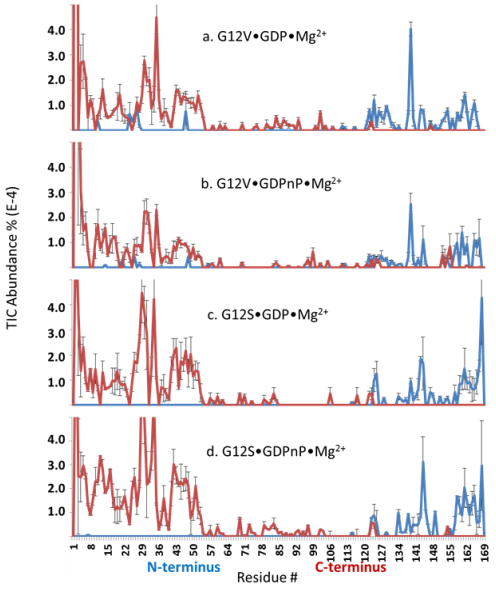

Figure 8. (a,b,c,d) Relative abundances of the holo (ligand-containing) fragment ions produced upon UVPD of G12V and G12S complexes, color coded as N-terminal holo fragments (blue traces) and C-terminal holo fragments (red traces). The abundances are plotted relative to the amino acid sequence along the x-axis to convey the relative efficiencies of backbone cleavages adjacent to each residue. Shaded regions (blue and lavender shading) highlight the regions of the most notable changes in UVPD fragmentation efficiencies between the GDP and GDPnP complexes and are expanded in Supplemental Figure 8.

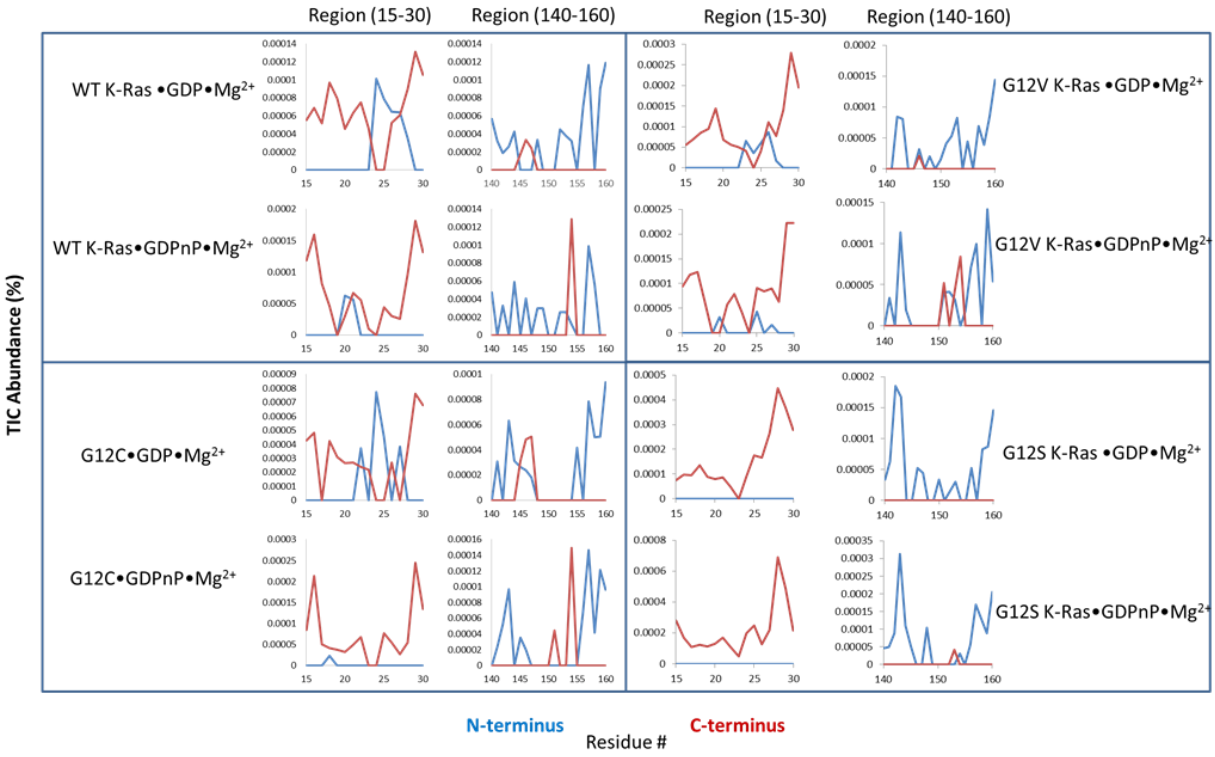

Figure 9. Expanded regions from Figure 1 and Supplemental Figure 7. Relative abundances of the holo-fragment ions produced upon UVPD of WT-K-Ras and G12C complexes, color coded as N-terminal fragments (blue traces) and C-terminal fragments (red traces).

To aid visualization of the regions demarcated by the holo fragment ions, the residues of particular interest are represented as space-filled models on the crystal structure of the protein, as illustrated for WT-K-Ras•GDP•Mg^{2+} complex in **Figure 10** for which the key regions are highlighted as purple spheres. In particular, backbone cleavages adjacent to the purple amino acids resulted in N-terminal and C-terminal fragment ions (primarily a- and x-type ions for UVPD) that retained the GDP•Mg^{2+} ligand. For the WT-K-Ras•GDP•Mg^{2+} complex, backbone cleavages at residues 4, 10-13, 26-28, 49, 112, 121, 123-125, and 145-147 resulted in overlapping holo fragment ions from both N and C terminal ions, and these residues are highlighted in **Figure 10**. The WT-K-

Ras•GDP•Mg^{2+}crystal structure (PDB: 4OBE) indicates that residues 13, 15-18, 30, 116-117, 119, 146 engage in electrostatic interactions with the GDP•Mg^{2+} ligand.

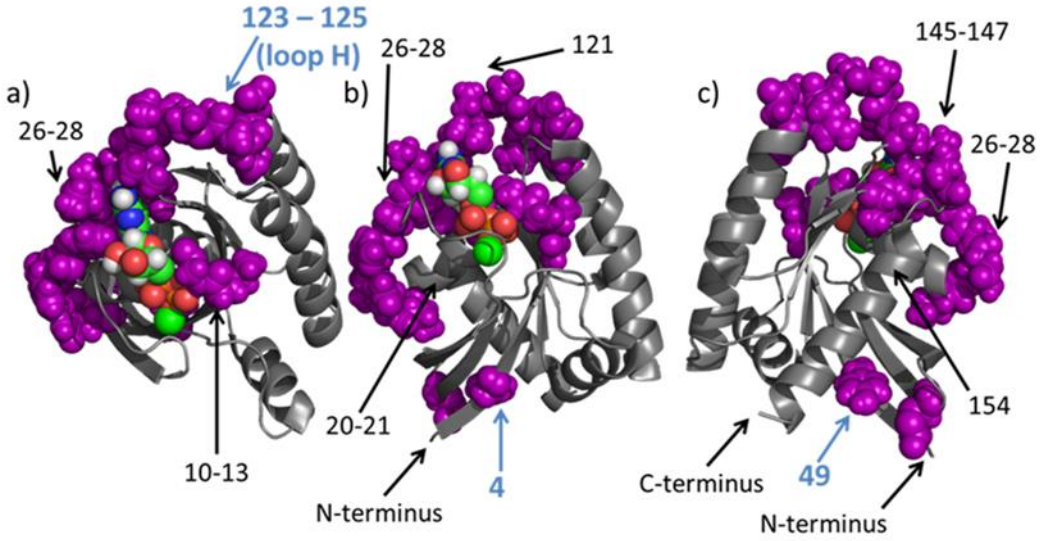

Figure 10. Space-filling models representing potential GDP•Mg^{2+} binding residues derived from the UVPD fragmentation patterns based on detection of holo (GDP-containing) fragment ions from the WT K-Ras•GDP•Mg^{2+} complexes (8+ charge state). The crystal structure represents the GDP-bound WT K-Ras construct (4OBE, WT K-Ras•GDP•Mg^{2+}) displayed as three views (a) top, (b) front side and (c) back side. The purple-colored residues are ones which are derived from the holo (GDP-containing) fragment ions. Specific amino acids are numbered in the crystal structures. Residues labelled in blue font indicate residues that exhibit potential ligand binding based on the UVPD results but do not appear to interact with the ligand in the 4OBE crystal structure.

Inspection of **Figures 7** and **10** indicate that the binding sites revealed from the UVPD results based on the formation of holo fragment ions parallel the predicted location of the ligand. Each amino acid lies within two to three residues of an amino acid known to engage in electrostatic interactions with the ligand, with the exception of three sites found from the UVPD data but not predicted from analysis of the crystal structure. These three cleavage sites that were indicative of K-Ras/GDP interactions based on the UVPD data

included residues 4, 49 and 123-125 (all labelled in blue font in **Figure 10**). One of these predicted sites (123-125) is particularly interesting. This extra site is found on loop H, which may be able to fold down and interact with the guanine portion of the GDP ligand, thus suggesting that the extended loop conformation modeled in the crystal structure may not reflect all of the conformations that play a functional role in ligand binding.

Evaluation of the holo fragment ions that differ for the K-Ras complexes containing GDP versus GDPnP offers insight into changes in the ligand binding mode as a function of the ligand identity. Two major changes are observed in the graphical displays shown **Figure 7** (comparison of **Fig. 7a** (WT K-Ras •GDP•Mg^{2+}) to **Fig. 7b** (WT K-Ras•GDPnP•Mg^{2+}), and comparison of **Fig. 7c** (G12C•GDP•Mg^{2+}) to **Fig. 7d** (G12C•GDPnP•Mg^{2+})). Similar plots are displayed in **Figures 8** and **9** for complexes containing the G12V and G12S variants. These regions of the UVPD data from residues 15 – 30 (shaded in blue in **Figure 7**) and 140 – 160 (shaded in lavender in **Figure 7**) are expanded in **Figure 9** for all variants. The first change is noted in the shift of N-terminal ions in the stretch of amino acids around 20 to 28 for both the WT K-Ras and G12C complexes. N-terminal holo ions arising from backbone cleavages adjacent to residues 24-28 are observed for the GDP complexes, but these N-terminal ions vanish for the GDPnP complexes and instead new N-terminal holo ions that originate from cleavages adjacent to residues 20-21 are observed. The second change is related to the formation of C-terminal holo ions in the region spanning residues 145 to 154 for the WT K-Ras, G12C and G12V complexes. C-terminal holo ions stemming from backbone cleavages adjacent to residues 145-147 are observed for the GDP complexes, but instead C-terminal holo ions that arise from cleavage adjacent to residue 154 are observed for the GDPnP complexes. We attribute these changes in the UVPD behavior to an alteration of the protein conformation that modulates the rigidity/flexibility of those local regions, as

consequently reflected in the backbone fragmentation efficiency. The implication is that the change in protein conformation arises from a shift in the ligand binding mode of both the phosphate portion and the guanine binding region in the WT K-Ras, G12C, and G12V complexes on going from GDP to GDPnP. For the complexes containing the G12S mutant, these same variations in the fragmentation patterns were not detected except for the formation of a single low abundance C-terminal holo ion arising from cleavage adjacent to residue 154. However, the complexes containing the G12S mutant display similar holo ion patterns (e.g. the overlap between N and C terminal holo ions seen between residues 123-125) that are consistently observed for all variants (**Figure 8**). Furthermore, when these two general regions (spanning residues 15-30 and 140-160) are superimposed on the crystal structures (4OBE for the K-Ras •GDP•Mg^{2+} complexes and 3GFT for the K-Ras •GDPnP•Mg^{2}+ complex), they appear to interact with one another (**Figure 7 e,f**), supporting the idea of a change in the binding motif upon substitution of GDPnP for GDP, one that is consistent across K-Ras variants.

6.5.2 Changes in UVPD fragmentation upon ligand exchange: Impact of GDP versus GDPnP

While analysis of the holo fragment ions upon UVPD reveals information about the ligand binding sites as described above, analysis of both the holo (with ligand) and apo (without ligand) fragment ions provides information about the efficiency of backbone cleavage across the protein, a factor that provides conformational information.[30,32] Based on past evidence, backbone cleavages are enhanced or suppressed upon UVPD depending on the flexibility/accessibility of different regions as well as whether those regions are shielded by a ligand (i.e. involved in binding interactions with the ligand) or engaged in other stabilizing intramolecular interactions. For the present study, we were particularly interested in evaluating the changes in backbone fragmentation relative to the impact of

the ligand: GDP versus GDPnP. Exchanging the diphosphate nucleotide (GDP) for the triphosphate surrogate (GDPnP) altered the efficiency of backbone cleavage between different residues, as shown in **Figure 11** for which the backbone cleavage efficiency upon UVPD is graphically displayed relative to each amino acid in the protein (summing fragment ions from both N-terminal and C-terminal cleavages on either side of each residue and including both holo and apo fragment ions).

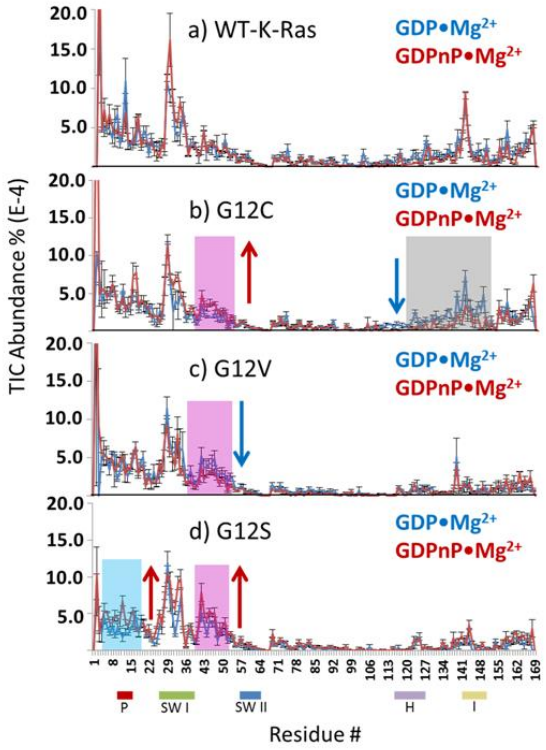

Figure 11. Relative ion abundances of the summed holo and apo fragment ions produced upon UVPD of complexes containing (a) WT K-Ras, (b) G12C, (c) G12V, or (d) G12S and GDP•Mg^{2+} (blue data trend) or GDPnP •Mg^{2+} (red data trend). The pink, gray, and blue shaded regions indicate regions of change specific to the G12X complexes in comparison to the WT K-Ras complexes. Upward or downward arrows adjacent to the shaded sections indicate those regions for which UVPD is enhanced (red) or suppressed (blue) upon ligand exchange from GDP to GDPnP. Relevant loops and switches are labelled underneath the x-axis using colors corresponding to Supplemental Figure 2. Calculated difference plots are shown in Figure 4.

Upon overlaying the UVPD trends for the GDP complexes (blue data trends in **Figure 11**) and GDPnP complexes (red data trends in **Figure 11**), many regions of the backbone show negligible changes in cleavage efficiency. The visualization of the similarities and differences are enhanced via construction of difference plots which are displayed in **Figure 12**.

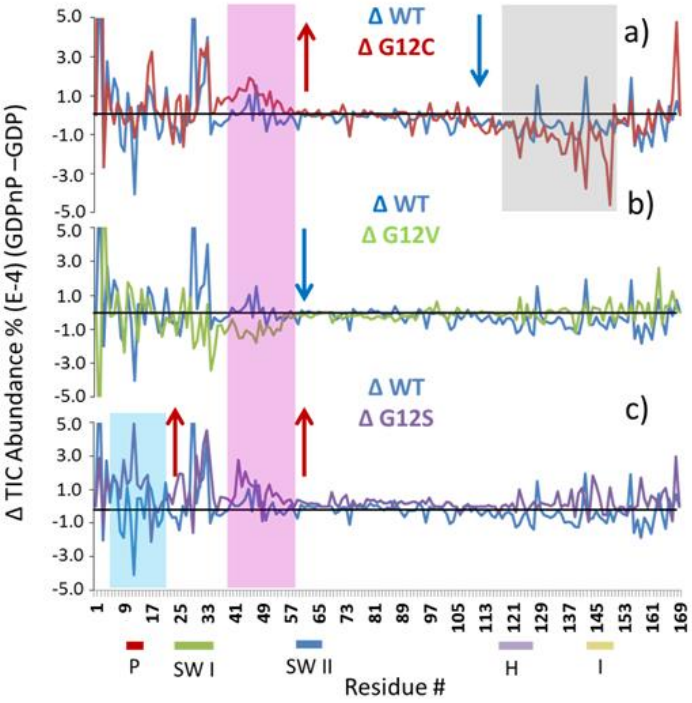

Figure 12. Plots showing the differences in abundances of fragment ions obtained from WT K-Ras and G12X [a) G12C, b) G12V, and c) G12S] complexes containing GDPnP or GDP (i.e. differences in abundances of fragment ions from cleavages along the backbone of the protein for each complementary pair of protein•GDP•Mg^{2+} and protein•GDPnP•Mg^{2+} complexes). The difference plots are constructed from the data shown in Figure 3. The difference plot for the WT K-Ras complexes (blue) is repeated in each of the three plots to facilitate comparisons to each mutant. Shaded regions relate directly to Figure 3. Upward or downward arrows adjacent to the shaded sections indicate those regions for which UVPD is enhanced or suppressed upon ligand exchange from GDP to GDPnP. Relevant loops and switches are labelled underneath the x-axis using colors corresponding to Supplemental Figure 2. Representative standard deviations are shown on the unprocessed plots in Figure 3.

There are small but reproducible variations in backbone cleavage efficiency upon UVPD for several regions with respect to the WT K-Ras—particularly related to amino acids 1, 3, 5-6, 9, 15-16, 20, 28, 30-32, 44, 46, 127, 141, and 154 for which backbone cleavage efficiencies increase for the GDPnP complexes relative to the GDP complexes and residues 2, 7-8, 11, 18, 24, 73, 102, 121-122, 125-126, 138-139, 147, 155, 157, 159, 160, and 164 for which backbone cleavages decrease for the GDPnP complexes relative to the GDP complexes.

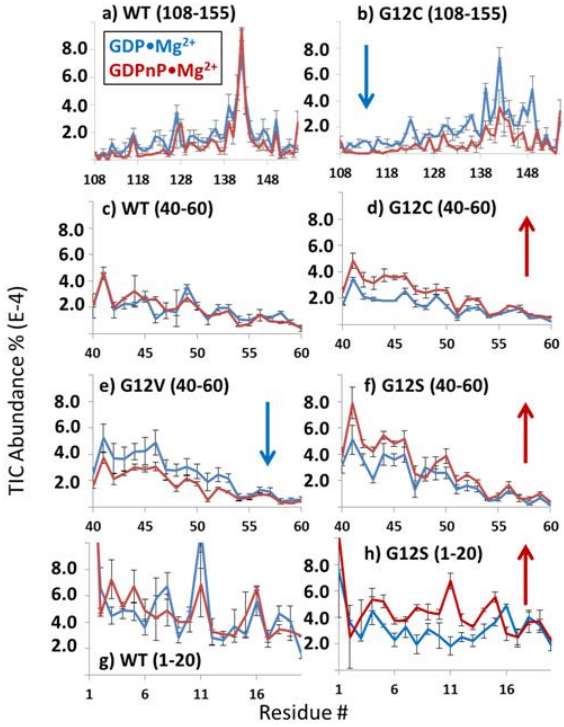

Figure 13. Expansions of UVPD fragmentation trends (from Figure 3) for specific regions along the backbone, particularly (a,b) residues 108-155 for WT K-Ras and the G12C mutant, (c-f) residues 40-60 for all four K-Ras proteins, and (g,h) residues 1-20 of WT K-Ras and the G12S mutant. The complete sets of UVPD fragmentation trends for summed holo and apo fragment ions are shown in Figure 3. The UVPD fragmentation trends are shown in blue for the GDP•Mg^{2+} complexes and red for the GDPnP•Mg^{2+} complexes . Upward or downward arrows indicate those regions for which UVPD is enhanced (red) or suppressed (blue) upon ligand exchange from GDP to GDPnP. (averaging over the entire section of amino acids).

The UVPD fragmentation trends are expanded for these regions in **Figure 13**, and the residues are highlighted on the structure of WT K-Ras in **Figure 14** for which blue-colored residues designate a reduction in backbone cleavage (suggesting ligand shielding or involvement in new intramolecular interactions) and red-colored residues denote an increase in backbone cleavage (suggesting greater flexibility or weakened intramolecular interactions) for the GDP versus GDPnP complexes. Considering all constructs analyzed, most of the enhancement in backbone cleavages for the GDPnP complexes occurs in the first 50 residues of K-Ras, whereas most of the suppression occurs beyond residue 100. We speculate that the observed variations reflect the interplay between the known guanine-binding residues (116,119,146), the ribose-binding residues (30,117) and the diphosphate-binding residues (13,15-18) of the complexes. Analysis of the UVPD trends in this way suggests that upon exchange of GDP for GDPnP, the guanine-binding region interacts more closely with GDPnP (tighter binding; decreased backbone cleavage upon UVPD), whereas the phosphate-binding and ribose-binding regions become more dynamic (weaker interactions with GDPnP; increased backbone cleavage upon UVPD). This hypothesis is further supported by the earlier observations that the abundances of the N-terminal holo fragment ions (i.e. ones associated with cleavages adjacent to residues 14-22) decreased significantly in the phosphate-binding region (**Figure 7**) going from the GDP to GDPnP complexes, suggesting reorientation of this stretch of the protein.

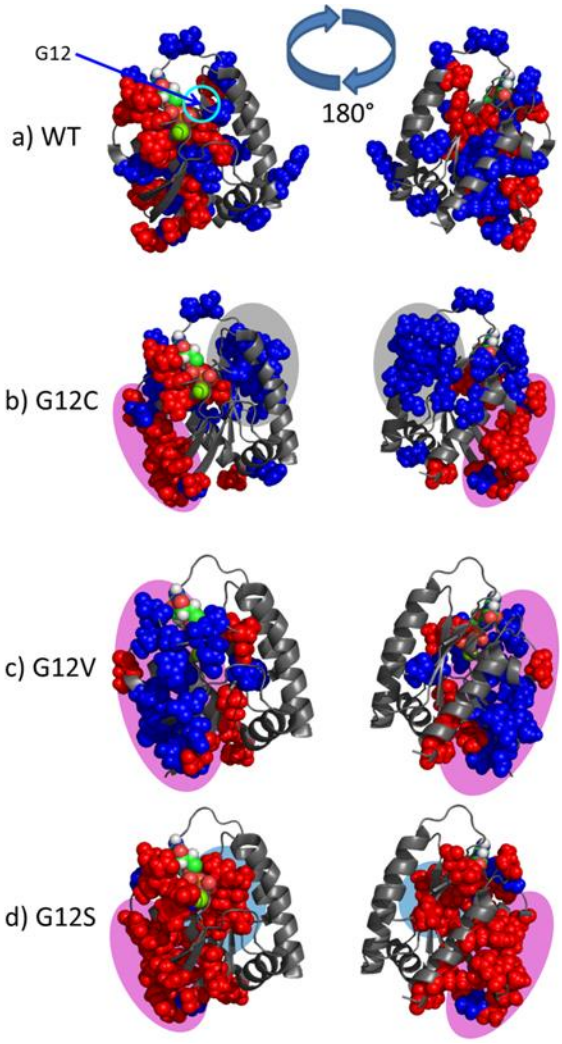

Figure 14. Space-filling models showing the suppression (blue residues) or enhancement (red residues) of backbone cleavages upon UVPD relative to ligand replacement (GDP versus GDPnP) for complexes containing (a) WT K-Ras, (b) G12C, (c) G12V, or (d) G12S constructs. The crystal structure represents the GDP-bound WT K-Ras construct from 4OBE. Regions shaded in light blue (residues 1-20), pink (residues 40-60) and gray (residues 120-150) correspond to the same regions shaded in Figure 3 which indicate regions of change specific to the G12X complexes. Cyan circle indicates the position of the G12 residue.

The primary regions of WT K-Ras that undergo the most significant changes in backbone cleavage efficiency upon UVPD of the GDPnP versus GDP complexes are

close to the ligand in the crystal structure, such as the P-loop (residues 9, 11, and 15-16), switch I (residues 30-32, ribose-binding stretch), and H loop (residues 121-122, and 125-126), for which backbone cleavages were suppressed in the P and H loops and enhanced in switch I. Additionally, there are a number of similarities in the backbone cleavage efficiencies for the GDP and GDPnP complexes for all K-Ras variants (with some exceptions), supporting that the unique changes induced by the G12X mutations are non-artifactual. The major similarities in the UVPD behavior of the GDP versus GDPnP complexes for the four proteins include an enhancement of backbone cleavage in the switch I region and suppression of backbone cleavages in loop H (116-125), loop I (143-150), the P-loop (10-14), and between residues 70-75. Each of these regions is known to interact directly with the nucleotide ligand with the exception of residues 70-75. From the crystal structures (4OBE and 3GFT in **Figure 7e,f**), R73 is on the edge of helix B which frays upon ligand exchange to GDPnP. UVPD fragmentation is enhanced in the region of switch I (residues 27-36) for the WT, G12C and G12S constructs but not the G12V construct, giving G12V an additional distinguishing feature. Additionally, the fact that residues on the H loop (116-125) of G12C exhibit a reduction in the UVPD backbone cleavage efficiency upon ligand exchange (**Figures 11b** and **12a** and **13b)** provides support that this loop interacts with the nucleotide in the guanine moiety in the G12C•GDP•Mg^{2+} complexes. Suppression of UVPD-induced backbone cleavages was less significant for the G12V complexes than for the WT or G12C complexes, whereas there was no change in the UVPD efficiency of G12S in in the H-loop region. Further evidence for the H loop/GDP interaction arises from the presence of holo ions containing residues 121 and 123-125, a result that was observed for all K-Ras variants.

6.5.3 UVPD of WT K-Ras versus G12X Complexes: Effects of G12C, G12V, and G12S mutations

As noted above, there are many similarities in the UVPD trends for the WT-K-Ras and G12X complexes upon exchange of GDP for GDPnP. However, there are also several distinctive differences in the UVPD behavior that specifically point to the impact of the G12X mutations, not the nature of the ligand. These differences are highlighted in **Figures 11** and **12** for three distinct regions of the proteins along with their general trend of enhanced or suppressed UVPD (upward or downward arrows), and the variations in backbone cleavage efficiency upon UVPD are expanded in **Figure 13**. The regions of particular interest are comprised of residues 5-20 (highlighted in blue), residues 40-60 (highlighted in pink) and residues 120-150 (highlighted in grey). Interestingly, the region spanning residues 40-60 displayed significant changes in UVPD fragmentation for the three mutants (G12C, G12V, G12S) compared to WT K-Ras (shaded in pink in **Figures 11** and **12,** and expanded in **13c-f**). UVPD fragmentation was enhanced across the backbone region covering residues 41-50 after ligand exchange of the G12C and G12S complexes. However, UVPD fragmentation was suppressed in this same general region for the G12V complexes. The G12C complexes uniquely exhibited suppression of backbone cleavages for a large portion of the C-terminus region (**Figure 14b,** residues 108-150; the stretch shaded in gray in **Figure 11b** and **Figure 12a** and expanded in **Figure 13b**) that was not observed for the WT-K-Ras or other variant complexes (again in each case comparing the GDP to GDPnP complexes). UVPD of the G12S complexes reveals a unique and significant enhancement in the P-loop (residues 10-14) (shaded in blue in **Figure 11d** and **Figure 12c** and expanded in **Figure 13h**) where the mutation has occurred. A table of all residues for which UVPD is enhanced or suppressed for each construct is included in **Table 1.**

K-Ras Construct	Enhanced (Res >+0.0001 TIC Abundance Change)	Suppressed (<-0.0001 TIC Abundance Change)
WT	1,3,5,6,9,15,16,20,28,30-32,44,46,127,141,154	2,7,8,11,18,24,73,102,121,122,125,126,138,139,147,155,157,159,160,164
G12C	1,3,15,16,24,26,30-32,41-48,50,167	2,11,22,25,73,92,121,122,125,127,128,130,132-136,138,140,141,144,145-148,155,158
G12V	2,3,6,9,10,13,25,162,167	1,8,12,19,23,28,29,31,33,34,37,41-46,48,51,53,138
G12S	1,3,5,6,8-15,23-25,29-33,41-47,50,140,143,144,148,157,164,167	2,16,28

Table 1. Summary of residues for which UVPD is enhanced or suppressed

A recent study compared the rates of hydrogen/deuterium exchange (HDX) for WT K-Ras and G12C complexes containing GDP or GDPnP.[47] While these experiments did not allow resolution of HDX rates of individual residues, it was determined that regions encompassing residues 7-19 and 114-120 exchanged at faster rates upon GDPnP ligation, an outcome suggestive of a more open conformation for the GDPnP complexes than the corresponding GDP-bound complexes.[47] These same two regions (7-19 and 114-120) also exhibit considerable changes in backbone cleavage efficiencies upon UVPD in the present study, thus further supporting that the tertiary structures of the proteins in solution are retained in the gas phase. In the same HDX study, one region of K-Ras underwent a slower rate of HDX upon GDPnP ligation (residues 38-52).[47] This region encapsulates the region noted in the present study for which there is a unique change in backbone cleavage upon UPVD for the G12X constructs (containing loop C and strand B). Although the change in the rate of HDX was small, the underpinning of the conformational structural change may be exaggerated upon transfer of the protein to the

gas phase from the solution phase or may signify a greater sensitivity of UVPD to small conformational changes.

The commonalties and differences in UVPD fragmentation trends with respect to the identity of the amino acid in position 12 (G,C,V,S) are striking. UVPD was enhanced in the region spanning residues 40 to 50 when position 12 was occupied by small polar amino acids that can form hydrogen bonds (C and S), whereas UVPD was suppressed when V, a bulky hydrophobic residue, was located at position 12. The fact that the presence of these single mutations alters the observed fragmentation behaviors differently reflects the exceptional sensitivity of UVPD to variations in protein structure and is consistent with the hypothesis that the identity of the particular substitution at the G12 position matters with respect to changes in protein structure or dynamics, with possible downstream functional effects that could impact the disease phenotype.

By using UVPD as a means to reveal subtle structural changes in the K-Ras complexes, the biological impact of mutation at residue 12 can be re-visited. As was first demonstrated in the homolog H-Ras (93% amino acid identity to K-Ras residues 1-166), mutation of the Gly12 site to any amino acid (except Pro) was found to produce an activated H-Ras.[49] The Gly12 site is situated within van der Waals distance of two catalytic residues that stimulate GTP hydrolysis in the complex between H-Ras and the GTPase-activating protein p120GAP, so bulkier substitutions at the Gly12 position would be expected to inhibit GTP hydrolysis, providing a molecular mechanism to explain H-Ras activation.[50] However, not all biological effects are readily explained by an increase in the proportion of Ras found in an activated state. For example, the particular identity of the amino acid substituted at the Gly12 site of H-Ras impacts the morphological phenotype of cells transformed with the corresponding genes.[49] Specifically for K-Ras, G12V and G12D variants were shown to generate distinct differences in the

(phospho)proteomic signatures of colorectal cancer cell lines.[51] Also, G12D substitutions in K-Ras showed activation of phosphatidylinositol 3-kinase (PI-3-K) and mitogen-activated protein extracellular signal-regulated kinase (MEK) signaling pathways, but G12C and G12V showed activated Ral signaling and decreased growth factor-dependent Akt activation, all in a non-small-cell lung carcinoma cell line.[52] In the latter case, molecular modeling suggested that the different sizes of the substitutions made at the Gly12 position may differently impact how the K-Ras variants interact with their downstream partners, in this case PI-3-K or itself (homodimerization), which subsequently impacts RaLGDS binding. Techniques to determine how mutations at the G12 position affect K-Ras conformation and flexibility would help to better elucidate the mechanisms behind differential binding affinity and selectivity. Towards these ends, molecular dynamics and potential mean force simulations of WT K-Ras and the G12D variant do suggest that this substitution increases flexibility in three regions: the P-loop (residues 10-14) switch-1 (residues 27-36) and switch-2 (residues 58-64) regions.[53] Notably, these regions are the same, or are in close proximity to sites where enhanced cleavage is observed (residues 15-16, 41-48) by the UVPD strategy presented herein.

The new evidence of structural variations of the K-Ras complexes containing single mutations based on the UVPD fragmentation allows further speculation about how each mutation might differently modulate the interactions of K-Ras with effector proteins (like Raf) as well as influence dimerization of K-Ras which is thought to be essential for interaction with Raf dimers for downstream signaling.[54,55] For instance, UVPD fragmentation was suppressed for the G12C•GDPnP•Mg^{2+} complex relative to the G12C•GDP•Mg^{2+} complex at the alpha helical interface region which involves alpha helices C and D (residues 86-105 and 126-138).[55] While dimers were not probed directly in the present study, the suppression of UVPD in alpha helices C and D for the

G12C•GDPnP•Mg^{2+} complex may indicate a stabilization of the alpha helical bundles in the monomer, potentially translating to a more stable alpha helical dimer which is the putative arrangement adopted for interaction with dimerized Raf.[55] In contrast, UVPD is suppressed in the effector region of the G12V•GDPnP•Mg^{2+} complex (relative to G12V•GDP•Mg^{2+}) which includes the beta sheet region (such as residues 41-45 of beta strand B) and Switch I (residues 27-36) which form the beta interface. While the beta interface dimer is not the ideal homodimer for downstream signaling, evidence for higher order oligomerization states (Ras nanoclusters), most likely mediated through multiple interfaces, has been observed and suggested to be an important determinant in signaling output and fidelity.[55,] For the third mutant probed (G12S), the increase in UVPD fragmentation at the P-loop (residues 8-15), switch I (residues 29-33), and beta strand B (residues 37-45) of the G12S•GDPnP•Mg^{2+} complex (relative to G12S•GDP•Mg^{2+}) suggests an increase in flexibility at the beta strand region in comparison to the G12V and G12C mutants. This overall increase in flexibility may serve a dual function to block effector binding and additionally to suppress the dimerization at the beta-interface which is the incorrect orientation for Raf signaling.[55] Although further studies will be required to fully characterize the impact of each mutant on K-Ras binding partners, the UVPD-MS technique described here clearly reveals the impact of single amino acid changes at the G12X position on more distant changes in protein conformation and flexibility.

6.6 CONCLUSION

The mutation of K-Ras at the twelfth residue from glycine to cysteine, serine, or valine was successfully probed by UVPD mass spectrometry based on comparisons of backbone fragmentation efficiencies of complexes containing GDP or GDPnP for all K-Ras protein variants. Overall, the changes in UVPD fragmentation efficiencies reflect

changes in protein conformation that are consistent with crystallographic or other experimental data. The changes in backbone cleavage efficiencies are attributed to variations in the flexibility or rigidity of the protein in specific regions, primarily due to variations in intra- or intermolecular interactions that are sensitive to single point mutations. A series of holo (ligand-containing) ions generated by UVPD offered a convenient means to map the GDP or GDPnP binding residues of K-Ras and also indicated that residues 121-125 interacted with the guanine portion of GDP or GDPnP. There was a notable increase in the UVPD backbone cleavage efficiency within the first 50 residues of K-Ras upon replacement of GDP by GDPnP, while at the same time there was a decrease in backbone cleavage efficiency in the region beyond the 100^{th} residue. These two regions are known to interact with the phosphate and guanine portions, respectively, suggesting a shift in the binding motif upon the ligand exchange. Most importantly, the series of G12X mutations resulted in unique changes in UVPD fragmentation of the K-Ras constructs that were significantly different upon ligand exchange. Interestingly, similarities were seen between hydrogen bonding amino acid substitutions (C,S) and the bulkier substitution (V). This supports that different K-Ras mutations contribute to specific downstream effects based on conformational or dynamic changes induced by the mutation. UVPD has proven to be novel structural tool for probing G12X K-Ras mutations and should be easily applied to other positional isomers of K-Ras.

6.7 REFERENCES

(1) Erichsen, H. C.; Chanock, S. J. *Br. J. Cancer* **2004**, *90* (4), 747–751.
(2) Rebbeck, T. R.; Ambrosone, C. B.; Bell, D. A.; Chanock, S. J.; Hayes, R. B.; Kadlubar, F. F.; Thomas, D. C. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13* (5), 681–687.
(3) Colicelli, J. *Sci. STKE Signal Transduct. Knowl. Environ.* **2004**, *2004* (250), RE13.
(4) Prior, I. A.; Lewis, P. D.; Mattos, C. *Cancer Res.* **2012**, *72* (10), 2457–2467.
(5) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(6) Miteva, Y. V.; Budayeva, H. G.; Cristea, I. M. *Anal. Chem.* **2013**, *85* (2), 749–768.
(7) Ngounou Wetie, A. G.; Sokolowska, I.; Woods, A. G.; Roy, U.; Loo, J. A.; Darie, C. C. *Proteomics* **2013**, *13* (3-4), 538–557.
(8) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(9) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.
(10) Fitzgerald, M. C.; West, G. M. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1193–1206.
(11) Engen, J. R. *Anal. Chem.* **2009**, *81* (19), 7870–7875.
(12) Konermann, L.; Pan, J.; Liu, Y.-H. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234.
(13) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(14) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(15) Ben-Nissan, G.; Sharon, M. *Chem. Soc. Rev.* **2011**, *40* (7), 3627–3637.
(16) Snijder, J.; Heck, A. J. R. *Annu. Rev. Anal. Chem.* **2014**, *7* (1), 43–64.
(17) Van Duijn, E. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 971–978.
(18) Vahidi, S.; Stocks, B. B.; Konermann, L. *Anal. Chem.* **2013**, *85* (21), 10471–10478.
(19) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2012**, *85* (1), 2–9.
(20) Schennach, M.; Breuker, K. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 164–168.
(21) Breuker, K.; Brüschweiler, S.; Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, *50* (4), 873–877.
(22) McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. *J. Am. Chem. Soc.* **2015**, *137* (39), 12667-76.
(23) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(24) Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.
(25) Jurneczko, E.; Barran, P. E. *The Analyst* **2011**, *136* (1), 20–28.
(26) Niu, S.; Rabuck, J. N.; Ruotolo, B. T. *Curr. Opin. Chem. Biol.* **2013**, *17* (5), 809–817.
(27) Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. *Nat. Methods* **2012**, *9* (11), 1084–1086.

(28) Cubrilovic, D.; Barylyuk, K.; Hofmann, D.; Walczak, M. J.; Gräber, M.; Berg, T.; Wider, G.; Zenobi, R. *Chem. Sci.* **2014**, *5* (7), 2794–2803.
(29) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(30) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128-9135.
(31) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. *Anal. Chem.* **2014**, *86* (1), 317–320.
(32) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(33) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(34) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.
(35) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(36) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(37) Zhang, H.; Cui, W.; Gross, M. L. *Int. J. Mass Spectrom.* **2013**, *354–355*, 288–291.
(38) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(39) Zhang, H.; Cui, W.; Gross, M. L.; Blankenship, R. E. *FEBS Lett.* **2013**, *587* (8), 1012–1020.
(40) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(41) Zhou, M.; Jones, C. M.; Wysocki, V. H. *Anal. Chem.* **2013**, *85* (17), 8262–8267.
(42) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(43) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(44) Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.; Brodbelt, J. *Anal. Chem.* **2014**, *86* (5), 2534–2542.
(45) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(46) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804-2812.
(47) Lim, S. M.; Westover, K. D.; Ficarro, S. B.; Harrison, R. A.; Choi, H. G.; Pacold, M. E.; Carrasco, M.; Hunter, J.; Kim, N. D.; Xie, T.; Sim, T.; Jänne, P. A.; Meyerson, M.; Marto, J. A.; Engen, J. R.; Gray, N. S. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 199–204.
(48) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. *Nature* **2013**, *503* (7477), 548–551.
(49) Seeburg, P. H.; Colby, W. W.; Capon, D. J.; Goeddel, D. V.; Levinson, A. D. *Nature* **1984**, *312* (5989), 71–75.

(50) Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmüller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A. *Science* **1997**, *277* (5324), 333–339.
(51) Hammond, D. E.; Mageean, C. J.; Rusilowicz, E. V.; Wickenden, J. A.; Clague, M. J.; Prior, I. A. *J. Proteome Res.* **2015**, *14* (3), 1535–1546.
(52) Ihle, N. T.; Byers, L. A.; Kim, E. S.; Saintigny, P.; Lee, J. J.; Blumenschein, G. R.; Tsao, A.; Liu, S.; Larsen, J. E.; Wang, J.; Diao, L.; Coombes, K. R.; Chen, L.; Zhang, S.; Abdelmelek, M. F.; Tang, X.; Papadimitrakopoulou, V.; Minna, J. D.; Lippman, S. M.; Hong, W. K.; Herbst, R. S.; Wistuba, I. I.; Heymach, J. V.; Powis, G. *J. Natl. Cancer Inst.* **2012**, *104* (3), 228–239.
(53) Chen, C.-C.; Er, T.-K.; Liu, Y.-Y.; Hwang, J.-K.; Barrio, M. J.; Rodrigo, M.; Garcia-Toro, E.; Herreros-Villanueva, M. *PLoS ONE* **2013**, *8* (2), e55793.
(54) Freeman, A. K.; Ritt, D. A.; Morrison, D. K. *Small GTPases* **2013**, *4* (3), 180–185.
(55) Muratcioglu, S.; Chavan, T. S.; Freed, B. C.; Jang, H.; Khavrutskii, L.; Freed, R. N.; Dyba, M. A.; Stefanisko, K.; Tarasov, S. G.; Gursoy, A.; Keskin, O.; Tarasova, N. I.; Gaponenko, V.; Nussinov, R. *Structure* **2015**, *23* (7), 1325–1335.

Chapter 7: Conclusion

Over the course of this dissertation two established methods, crosslinking and native mass spectrometry, were improved upon through the incorporation of ultraviolet photodissociation (UVPD), an alternative ion activation technique. While the use of mass spectrometry to elucidate three-dimensional biomolecule representation remains elusive, the techniques that were detailed have the ability to yield distance restraints of residues from intra- and inter-protein crosslinking from a top-down perspective as well as use native-mass spectrometry to gain detailed information about the location and nature of conformational changes of active proteins.

While the use of top-down MS for analysis of crosslinked proteins has been previously undertaken in other labs, the results were underwhelming. For example, the previous studies revealed only two intra-protein crosslinks for the small protein ubiquitin. This was due to the lack of separation of the isobaric crosslinked species. Without separation the crosslinks are convoluted from the fragmentation resulting in a loss of site-specific crosslinking information. In the method presented in this dissertation separation of crosslinked ubiquitin and hexameric insulin was performed using reversed phase chromatography leading to an overall increase, to four, in crosslink identifications for ubiquitin and the first example of top-down inter-protein crosslink elucidation of insulin. The inter-protein crosslink elucidation resulted in three total identifications. UVPD was also key in identification of said crosslink species whereas collisional activation fell short due to lack of sequence coverage as well as undesirable internal fragmentation.

Native mass spectrometry has been a very attractive technique for analyzing biomolecules and larger order structures. However, little had been done with regards to determining conformational changes of proteins either from ligand exchange, enzymatic processes, or amino acid mutations. Throughout Chapters 4-6 these types of specific conformational changes were probed by UVPD with a twofold interest. Firstly, to determine how sensitive UVPD was to the conformational changes of proteins and secondly, to glean biological insights that may not be readily available through crystallographic or NMR studies.

It was shown in Chapter 4 that UVPD successfully can decipher loop-movements of DHFR upon inhibition by methotrexate. Upon this realization the method was extended to study clinically relevant DHFR mutations from TMP resistant *E. Coli* bacterium. During this analysis size exclusion chromatography was also employed to quantify the relative k_{off} rates of each mutant with respect to the substrate, co-factor, and its known inhibitors. In the end it was found that by the combined use of UVPD, SEC, and kinetic measurements the P21L and W30R mutations promoted TMP resistance through two separate mechanisms. A novel propargyl-linked antifolate was also found to have potent inhibition of the P21L mutant in comparison to the WT DHFR.

Finally, in Chapter 6 the oncogene K-Ras was studied. Briefly, UVPD mass spectra were collected for both the on and off states of the protein, meaning GDP and GTP (GDPnP) bound, respectively. These UVPD spectra were compared to understand the conformational changes of the WT protein. Several mutants (G12C, G12S, G12V)

were also compared in their on and off states to the WT protein to delineate where the mutations caused a difference structure between the two states. It was found that each of the mutants effected the proteins' conformational changes, as detected by UVPD, in specific unique regions of the protein. This is suggestive that each of these mutants may be driving cancer proliferation through different mechanisms of disease and should be treated as unique biological problems to be solved in cancer biology.

Through these studies it has been illustrated that native-MS UVPD can be used to probe conformational changes due to amino acid mutations and ligand exchange or inhibition of proteins in addition to the standard advantages of native-MS for protein analysis. These standard advantages include the ability to monitor topology and oligomerization of protein complexes as well as to map ligand binding to calculate dissociation constants. With these compounded advantages it is now feasible to consider applying this method in a high-throughput manner by use of a nano-spray robot such as an Advion Nanomate. The objective would be to screen potential inhibitors against an assortment of clinical oncomutations to quickly indicate if specific mutations have a preference for particular inhibitors over other lead inhibitors. UVPD would be used in the same manner as described in Chapter 6 to determine which inhibitors effect the proteins' conformational space and the location of the change. This would be a competitive screening method due to the low quantities of proteins and inhibitors as well as data analysis that has the potential to be completely automated.

Additionally, SEC was found to very useful in comparing the stabilities of protein-ligand complexes and should be continued to be used to probe protein-ligand

dissociation. It would be useful to develop the analytical method to be quantifiable. It would also be interesting to study if ion exchange chromatography would be sensitive enough to determine conformational changes based on the protein or protein-complexes elution as the concentration of salt is increased while it is interfaced to the mass spectrometer.

References

Chapter 1

(1) Kendrew, J. C.; Parrish, R. G. *Proc. R. Soc. Lond. Math. Phys. Eng. Sci.* **1957**, *238* (1214), 305–324.

(2) Shi, Y. *Cell* **2014**, *159* (5), 995–1014.

(3) Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. *FEBS J.* **2008**, *275* (1), 1–21.

(4) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci.* **2007**, *104* (23), 9615–9620.

(5) Wüthrich, K. *Nat. Struct. Mol. Biol.* **2001**, *8* (11), 923–925.

(6) Søndergaard, C. R.; Garrett, A. E.; Carstensen, T.; Pollastri, G.; Nielsen, J. E. *J. Med. Chem.* **2009**, *52* (18), 5673–5684.

(7) Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Mol. Biol.* **2002**, *319* (1), 209–227.

(8) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.

(9) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682.

(10) Rappsilber, J. *J. Struct. Biol.* **2011**, *173* (3), 530–540.

(11) Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. *Nat. Methods* **2015**, *12* (12), 1179–1184.

(12) Schmidt, C.; Robinson, C. V. *Nat. Protoc.* **2014**, *9* (9), 2224–2236.

(13) Suckau, D.; Mak, M.; Przybylski, M. *Proc. Natl. Acad. Sci.* **1992**, *89* (12), 5630–5634.

(14) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (15), 7391–7397.

(15) Zhou, Y.; Vachet, R. W. *Anal. Chem.* **2013**, *85* (20), 9664–9670.

(16) Zhang, H.; Liu, H.; Blankenship, R. E.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2015**, *27* (1), 178–181.

(17) Takahashi, K. *J. Biol. Chem.* **1968**, *243* (23), 6171–6179.

(18) Fliss, H.; Viswanatha, T. *Can. J. Biochem.* **1979**, *57* (11), 1267–1272.

(19) Yem, A. W.; Epps, D. E.; Mathews, W. R.; Guido, D. M.; Richard, K. A.; Staite, N. D.; Deibel, M. R. *J. Biol. Chem.* **1992**, *267* (5), 3122–3128.

(20) Zhou, Y.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (4), 708–717.

(21) Houston, L. L.; Walsh, K. A. *Biochemistry (Mosc.)* **1970**, *9* (1), 156–166.

(22) Horton, H. R.; Koshland, D. E. *J. Am. Chem. Soc.* **1965**, *87*, 1126–1132.

(23) Zhou, Y.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (5), 899–907.

(24) Madsen, J. A.; Yin, Y.; Qiao, J.; Gill, V.; Renganathan, K.; Fu, W.-Y.; Smith, S.; Anderson, J. *Anal. Chem.* **2016**, *88* (4), 2478–2488.

(25) Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C. *Anal. Chem.* **2012**, *84* (10), 4411–4418.

(26) Zhang, B.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2015**, *27* (3), 552–555.

(27) Gross, D. S.; Williams, E. R. *Int. J. Mass Spectrom. Ion Process.* **1996**, *157-158*, 305–318.

(28) Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L. *Anal. Chem.* **2011**, *83* (1), 311–318.
(29) Hambly, D.; Gross, M. *Int. J. Mass Spectrom.* **2007**, *259* (1–3), 124–129.
(30) Jones, L. M.; Sperry, J.; Carroll, J.; Gross, M. L. *Anal. Chem.* **2011**, *83* (20), 7657–7661.
(31) Hochleitner, E. O.; Borchers, C.; Parker, C.; Bienstock, R. J.; Tomer, K. B. *Protein Sci.* **2000**, *9* (3), 487–496.
(32) Kahraman, A.; Herzog, F.; Leitner, A.; Rosenberger, G.; Aebersold, R.; Malmström, L. *PLOS ONE* **2013**, *8* (9), e73411.
(33) Zybailov, B. L.; Glazko, G. V.; Jaiswal, M.; Raney, K. D. *J. Proteomics Bioinform.* **2013**, *6* (Suppl 2), 001.
(34) Chait, B. T.; Cadene, M.; Olinares, P. D.; Rout, M. P.; Shi, Y. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (6), 952–965.
(35) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. *Mol. Cell. Proteomics MCP* **2010**, *9* (8), 1634–1649.
(36) Merkley, E. D.; Rysavy, S.; Kahraman, A.; Hafen, R. P.; Daggett, V.; Adkins, J. N. *Protein Sci.* **2014**, *23* (6), 747–759.
(37) Pham, N. D.; Parker, R. B.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2013**, *17* (1), 90–101.
(38) Gomes, A. F.; Gozzo, F. C. *J. Mass Spectrom. JMS* **2010**, *45* (8), 892–899.
(39) Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. *Mol. Cell. Proteomics* **2011**, *10* (1), M110.002212.
(40) Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyö, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125* (9), 2416–2425.
(41) Engen, J. R. *Anal. Chem.* **2009**, *81* (19), 7870–7875.
(42) Baldwin, R. L. *Proteins Struct. Funct. Bioinforma.* **2011**, *79* (7), 2021–2026.
(43) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(44) McAllister, R. G.; Konermann, L. *Biochemistry (Mosc.)* **2015**, *54* (16), 2683–2692.
(45) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(46) Hayashi, N.; Kuyama, H.; Nakajima, C.; Kawahara, K.; Miyagi, M.; Nishimura, O.; Matsuo, H.; Nakazawa, T. *Biochemistry (Mosc.)* **2014**, *53* (11), 1818–1826.
(47) Cummins, D. J.; Espada, A.; Novick, S. J.; Molina-Martin, M.; Stites, R. E.; Espinosa, J. F.; Broughton, H.; Goswami, D.; Pascal, B. D.; Dodge, J. A.; Chalmers, M. J.; Griffin, P. R. *Anal. Chem.* **2016**, *88* (12), 6607–6614.
(48) Abzalimov, R. R.; Kaltashov, I. A. *Anal. Chem.* **2010**, *82* (3), 942–950.
(49) Hoerner, J. K.; Xiao, H.; Dobo, A.; Kaltashov, I. A. *J. Am. Chem. Soc.* **2004**, *126* (24), 7709–7717.
(50) Bobst, C. E.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (11), 5225–5231.
(51) Abzalimov, R. R.; Kaplan, D. A.; Easterling, M. L.; Kaltashov, I. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1514–1517.

(52) Wang, H.; Shu, Q.; Rempel, D. L.; Frieden, C.; Gross, M. L. *Biochemistry (Mosc.)* **2015**, *54* (42), 6475–6481.
(53) Zhang, Y.; Rempel, D. L.; Zhang, J.; Sharma, A. K.; Mirica, L. M.; Gross, M. L. *Proc. Natl. Acad. Sci.* **2013**, *110* (36), 14604–14609.
(54) Malito, E.; Faleri, A.; Surdo, P. L.; Veggi, D.; Maruggi, G.; Grassi, E.; Cartocci, E.; Bertoldi, I.; Genovese, A.; Santini, L.; Romagnoli, G.; Borgogni, E.; Brier, S.; Passo, C. L.; Domina, M.; Castellino, F.; Felici, F.; Veen, S. van der; Johnson, S.; Lea, S. M.; Tang, C. M.; Pizza, M.; Savino, S.; Norais, N.; Rappuoli, R.; Bottomley, M. J.; Masignani, V. *Proc. Natl. Acad. Sci.* **2013**, *110* (9), 3304–3309.
(55) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(56) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(57) Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683–693.
(58) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2012**, *85* (1), 2–9.
(59) Skinner, O. S.; McLafferty, F. W.; Breuker, K. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (6), 1011–1014.
(60) Schennach, M.; Breuker, K. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 164–168.
(61) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(62) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. *Nature* **2014**, *510* (7503), 172–175.
(63) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3* (2), 126–132.
(64) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(65) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16* (1), 1–23.
(66) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(67) Jurneczko, E.; Barran, P. E. *The Analyst* **2011**, *136* (1), 20–28.
(68) Niu, S.; Rabuck, J. N.; Ruotolo, B. T. *Curr. Opin. Chem. Biol.* **2013**, *17* (5), 809–817.
(69) Snijder, J.; Heck, A. J. R. *Annu. Rev. Anal. Chem.* **2014**, *7* (1), 43–64.
(70) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(71) Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. *Nat. Commun.* **2013**, *4*, 1985.
(72) Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 190–208.
(73) Zhou, M.; Wysocki, V. H. *Acc. Chem. Res.* **2014**, *47* (4), 1010–1018.
(74) Song, Y.; Nelp, M. T.; Bandarian, V.; Wysocki, V. H. *ACS Cent. Sci.* **2015**, *1* (9), 477–487.
(75) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.

(76) Hall, Z.; Politis, A.; Robinson, C. V. *Structure* **2012**, *20* (9), 1596–1609.
(77) Cong, X.; Liu, Y.; Liu, W.; Liang, X.; Russell, D. H.; Laganowsky, A. *J. Am. Chem. Soc.* **2016**, *138* (13), 4346–4349.
(78) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.
(79) Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V. *Nat. Methods* **2016**, *13* (4), 333–336.
(80) Bechara, C.; Nöll, A.; Morgner, N.; Degiacomi, M. T.; Tampé, R.; Robinson, C. V. *Nat. Chem.* **2015**, *7* (3), 255–262.
(81) McAlary, L.; Yerbury, J. J.; Aquilina, J. A. *Sci. Rep.* **2013**, *3*.
(82) Cubrilovic, D.; Haap, W.; Barylyuk, K.; Ruf, A.; Badertscher, M.; Gubler, M.; Tetaz, T.; Joseph, C.; Benz, J.; Zenobi, R. *ACS Chem. Biol.* **2014**, *9* (1), 218–226.
(83) Cubrilovic, D.; Barylyuk, K.; Hofmann, D.; Walczak, M. J.; Gräber, M.; Berg, T.; Wider, G.; Zenobi, R. *Chem. Sci.* **2014**, *5* (7), 2794–2803.
(84) Yin, S.; Loo, J. A. *Int. J. Mass Spectrom.* **2011**, *300* (2-3), 118–122.
(85) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(86) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2010**, *132* (25), 8602–8609.
(87) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(88) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(89) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(90) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72* (3), 563–573.
(91) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(92) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(93) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(94) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(95) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(96) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804–2812.
(97) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.

(98) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(99) Sawaya, M. R.; Kraut, J. *Biochemistry (Mosc.)* **1997**, *36* (3), 586–603.

Chapter 2

(1) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

Chapter 3

(1) A. Sinz, *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
(2) A. D. Catherman, O. S. Skinner, N. L. Kelleher, *Biochem. Biophys. Res. Commun.* **2014**, *445*, 683–693.
(3) H. Zhou, Z. Ning, A. E. Starr, M. Abu-Farha, D. Figeys, *Anal. Chem.* **2012**, *84*, 720–734.
(4) J. S. Brodbelt, *Chem. Soc. Rev.* **2014**, *43*, 2757–2783.
(5) A. W. Jones, H. J. Cooper, *Analyst* **2011**, *136*, 3419–3429.
(6) M. E. Belov, E. Damoc, E. Denisov, P. D. Compton, S. Horning, A. A. Makarov, N. L. Kelleher, *Anal. Chem.* **2013**, *85*, 11163–11173.
(7) J. B. Shaw, W. Li, D. D. Holden, Y. Zhang, J. Griep-Raming, R. T. Fellers, B. P. Early, P. M. Thomas, N. L. Kelleher, J. S. Brodbelt, *J. Am. Chem. Soc.* **2013**, *135*, 12646–12651.
(8) A. D. Catherman, K. R. Durbin, D. R. Ahlf, B. P. Early, R. T. Fellers, J. C. Tran, P. M. Thomas, N. L. Kelleher, *Mol. Cell. Proteomics* **2013**, *12*, 3465–3473.
(9) G. H. Kruppa, J. Schoeniger, M. M. Young, *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155–162.
(10) I. K. Webb, M. Mentinova, W. M. McGee, S. A. McLuckey, *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 733–743.
(11) E. D. Merkley, S. Rysavy, A. Kahraman, R. P. Hafen, V. Daggett, J. N. Adkins, *Protein Sci.* **2014**, *23*, 747–759.

Chapter 4

(1) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(2) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(3) Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.; Brodbelt, J. *Anal. Chem.* **2014**, *86* (5), 2534–2542.
(4) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. *J. Am. Chem. Soc.* **2009**, *131* (35), 12801–12808.
(5) Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. *J. Mass Spectrom.* **2004**, *39* (3), 322–328.

(6) Wang, G.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (15), 7293–7298.
(7) Stefanowicz, P.; Kijewska, M.; Szewczuk, Z. *Anal. Chem.* **2014**, *86* (15), 7247–7251.
(8) Bobst, C. E.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (11), 5225–5231.
(9) Nagy, K.; Redeuil, K.; Rezzi, S. *Anal. Chem.* **2009**, *81* (22), 9365–9371.
(10) Robinson, E. W.; Leib, R. D.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (10), 1470–1480.
(11) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(12) Hall, Z.; Hernández, H.; Marsh, J. A.; Teichmann, S. A.; Robinson, C. V. *Structure* **2013**, *21* (8), 1325–1337.
(13) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.
(14) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. *J. Am. Chem. Soc.* **2008**, *130* (35), 11574–11575.
(15) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(16) Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.
(17) Simmons, D. A.; Dunn, S. D.; Konermann, L. *Biochemistry (Mosc.)* **2003**, *42* (19), 5896–5905.
(18) Wang, F.; Tang, X. *Biochemistry (Mosc.)* **1996**, *35* (13), 4069–4078.
(19) Wright, P. J.; Zhang, J.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1906–1913.
(20) Zhou, M.; Jones, C. M.; Wysocki, V. H. *Anal. Chem.* **2013**, *85* (17), 8262–8267.
(21) Breuker, K.; Brüschweiler, S.; Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, *50* (4), 873–877.
(22) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(23) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2010**, *132* (25), 8602–8609.
(24) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(25) Modzel, M.; Stefanowicz, P.; Szewczuk, Z. *Rapid Commun. Mass Spectrom.* **2012**, *26* (23), 2739–2744.
(26) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(27) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(28) Catherman, A. D.; Durbin, K. R.; Ahlf, D. R.; Early, B. P.; Fellers, R. T.; Tran, J. C.; Thomas, P. M.; Kelleher, N. L. *Mol. Cell. Proteomics* **2013**, *12* (12), 3465–3473.
(29) Skinner, O. S.; Catherman, A. D.; Early, B. P.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L. *Anal. Chem.* **2014**.
(30) Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L. *Anal. Chem.* **2013**, *85* (23), 11163–11173.
(31) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(32) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(33) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.

(34) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(35) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(36) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, 1–9.
(37) Zhang, H.; Cui, W.; Gross, M. L.; Blankenship, R. E. *FEBS Lett.* **2013**, *587* (8), 1012–1020.
(38) Cui, W.; Rohrs, H. W.; Gross, M. L. *Analyst* **2011**, *136* (19), 3854–3864.
(39) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308-10314.
(40) Vahidi, S.; Stocks, B. B.; Konermann, L. *Anal. Chem.* **2013**, *85* (21), 10471–10478.
(41) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(42) Zhang, H.; Cui, W.; Gross, M. L. *Int. J. Mass Spectrom.* **2013**, *354–355*, 288–291.
(43) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.
(44) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(45) Clarke, D. J.; Murray, E.; Hupp, T.; Mackay, C. L.; Langridge-Smith, P. R. R. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (8), 1432–1440.
(46) Zhang, H.; Cui, W.; Wen, J.; Blankenship, R. E.; Gross, M. L. *Anal. Chem.* **2011**, *83* (14), 5598–5606.
(47) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. *Anal. Chem.* **2014**, *86* (1), 317–320.
(48) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(49) Harvey, S. R.; Porrini, M.; Konijnenberg, A.; Clarke, D. J.; Tyler, R. C.; Langridge-Smith, P. R. R.; MacPhee, C. E.; Volkman, B. F.; Barran, P. E. *J. Phys. Chem. B* **2014**, *118* (43), 12348–12359.
(50) Huennekens, F. M. *Adv. Enzyme Regul.* **1994**, *34*, 397–419.
(51) Sawaya, M. R.; Kraut, J. *Biochemistry (Mosc.)* **1997**, *36* (3), 586–603.
(52) Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119–140.
(53) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(54) Wang, Z.; Singh, P.; Czekster, C. M.; Kohen, A.; Schramm, V. L. *J. Am. Chem. Soc.* **2014**, *136* (23), 8333–8341.
(55) Wan, Q.; Bennett, B. C.; Wilson, M. A.; Kovalevsky, A.; Langan, P.; Howell, E. E.; Dealwis, C. *Proc. Natl. Acad. Sci.* **2014**, *111* (51), 18225–18230.
(56) Liu, C. T.; Francis, K.; Layfield, J. P.; Huang, X.; Hammes-Schiffer, S.; Kohen, A.; Benkovic, S. J. *Proc. Natl. Acad. Sci.* **2014**, *111* (51), 18231–18236.
(57) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. *Acc. Chem. Res.* **2015**.
(58) Liu, C. T.; Layfield, J. P.; Stewart, R. J.; French, J. B.; Hanoian, P.; Asbury, J. B.; Hammes-Schiffer, S.; Benkovic, S. J. *J. Am. Chem. Soc.* **2014**, *136* (29), 10349–10360.
(59) Osborne, M. J.; Schnell, J.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. *Biochemistry (Mosc.)* **2001**, *40* (33), 9846–9859.
(60) Hatse, S.; De Clercq, E.; Balzarini, J. *Biochem. Pharmacol.* **1999**, *58* (4), 539–555.

(61) Rajagopalan, P. T. R.; Zhang, Z.; McCourt, L.; Dwyer, M.; Benkovic, S. J.; Hammes, G. G. *Proc. Natl. Acad. Sci.* **2002**, *99* (21), 13481–13486.
(62) Li, R.; Sirawaraporn, R.; Chitnumsub, P.; Sirawaraporn, W.; Wooden, J.; Athappilly, F.; Turley, S.; Hol, W. G. J. *J. Mol. Biol.* **2000**, *295* (2), 307–323.
(63) Brunet, C.; Antoine, R.; Dugourd, P.; Canon, F.; Giuliani, A.; Nahon, L. *J. Chem. Phys.* **2013**, *138* (6), 064301.
(64) Tsai, C. S. *Biomacromolecules: Introduction to structure, function and informatics*; John Wiley & Sons, 2007.

Chapter 5

(1) Cullen, I. M.; Manecksha, R. P.; McCullagh, E.; Ahmad, S.; O'Kelly, F.; Flynn, R.; McDermott, T. E. D.; Murphy, P.; Grainger, R.; Fennell, J. P.; Thornhill, J. A. *Ir. J. Med. Sci.* **2012**, *182* (1), 81–89.
(2) Duffy, M. A.; Hernandez-Santiago, V.; Orange, G.; Davey, P. G.; Guthrie, B. *Br J Gen Pr.* **2013**, *63* (609), e238–e243.
(3) Pallett, A.; Hand, K. *J. Antimicrob. Chemother.* **2010**, *65* (suppl 3), iii25–iii33.
(4) Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119–140.
(5) Huovinen, P. *Antimicrob. Agents Chemother.* **1987**, *31* (10), 1451–1456.
(6) Toprak, E.; Veres, A.; Michel, J.-B.; Chait, R.; Hartl, D. L.; Kishony, R. *Nat. Genet.* **2012**, *44* (1), 101–105.
(7) Watson, M.; Liu, J.-W.; Ollis, D. *FEBS J.* **2007**, *274* (10), 2661–2671.
(8) Bryce, A.; Hay, A. D.; Lane, I. F.; Thornton, H. V.; Wootton, M.; Costelloe, C. *BMJ* **2016**, *352*, i939.
(9) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.
(10) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(11) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682.
(12) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(13) Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. *Nat. Commun.* **2013**, *4*, 1985.
(14) Zhou, M.; Wysocki, V. H. *Acc. Chem. Res.* **2014**, *47* (4), 1010–1018.
(15) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(16) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128–9135.
(17) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. *Nature* **2014**, *510* (7503), 172–175.
(18) Snijder, J.; van de Waterbeemd, M.; Damoc, E.; Denisov, E.; Grinfeld, D.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A.; Heck, A. J. R. *J. Am. Chem. Soc.* **2014**, *136* (20), 7295–7299.
(19) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3* (2), 126–132.
(20) Cong, X.; Liu, Y.; Liu, W.; Liang, X.; Russell, D. H.; Laganowsky, A. *J. Am. Chem. Soc.* **2016**, *138* (13), 4346–4349.
(21) Konermann, L.; Pan, J.; Liu, Y.-H. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234.

(22) Ben-Nissan, G.; Sharon, M. *Chem. Soc. Rev.* **2011**, *40* (7), 3627–3637.
(23) Holding, A. N. *Methods* **2015**, *89*, 54–63.
(24) Stefanowicz, P.; Kijewska, M.; Szewczuk, Z. *Anal. Chem.* **2014**, *86* (15), 7247–7251.
(25) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(26) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(27) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(28) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(29) Canon, F.; Milosavljević, A. R.; van der Rest, G.; Réfrégiers, M.; Nahon, L.; Sarni-Manchado, P.; Cheynier, V.; Giuliani, A. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8377–8381.
(30) Yin, S.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 899–907.
(31) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(32) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804–2812.
(33) Morrison, L. J.; Brodbelt, J. S. *Analyst* **2015**, *141* (1), 166–176.
(34) A. Loo, J.; A. Benchaar, S.; Zhang, J. *Mass Spectrom.* **2013**, *2* (Spec Iss).
(35) Bao, J.; Krylova, S. M.; Cherney, L. T.; Blanc, J. C. Y. L.; Pribil, P.; Johnson, P. E.; Wilson, D. J.; Krylov, S. N. *Analyst* **2015**, *140* (4), 990–994.
(36) Cherney, L. T.; Krylov, S. N. *Anal. Chem.* **2016**, *88* (7), 4063–4070.
(37) Muneeruddin, K.; Thomas, J. J.; Salinas, P. A.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (21), 10692–10699.
(38) Wang, G.; Kaltashov, I. A. *Anal. Chem.* **2014**, *86* (15), 7293–7298.
(39) Flensburg, J.; Sköld, O. *Eur. J. Biochem.* **1987**, *162* (3), 473–476.
(40) Lombardo, M. N.; G-Dayanandan, N.; Wright, D. L.; Anderson, A. C. *ACS Infect. Dis.* **2016**, *2* (2), 149–156.
(41) G-Dayanandan, N.; Paulsen, J. L.; Viswanathan, K.; Keshipeddy, S.; Lombardo, M. N.; Zhou, W.; Lamb, K. M.; Sochia, A. E.; Alverson, J. B.; Priestley, N. D.; Wright, D. L.; Anderson, A. C. *J. Med. Chem.* **2014**, *57* (6), 2643–2656.
(42) Keshipeddy, S.; Reeve, S. M.; Anderson, A. C.; Wright, D. L. *J. Am. Chem. Soc.* **2015**, *137* (28), 8983–8990.
(43) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(44) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.
(45) Kopytek, S. J.; Dyer, J. C. D.; Knapp, G. S.; Hu, J. C. *Antimicrob. Agents Chemother.* **2000**, *44* (11), 3210–3212.

Chapter 6

(1) Erichsen, H. C.; Chanock, S. J. *Br. J. Cancer* **2004**, *90* (4), 747–751.

(2) Rebbeck, T. R.; Ambrosone, C. B.; Bell, D. A.; Chanock, S. J.; Hayes, R. B.; Kadlubar, F. F.; Thomas, D. C. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13* (5), 681–687.
(3) Colicelli, J. *Sci. STKE Signal Transduct. Knowl. Environ.* **2004**, *2004* (250), RE13.
(4) Prior, I. A.; Lewis, P. D.; Mattos, C. *Cancer Res.* **2012**, *72* (10), 2457–2467.
(5) Konermann, L.; Vahidi, S.; Sowole, M. A. *Anal. Chem.* **2014**, *86* (1), 213–232.
(6) Miteva, Y. V.; Budayeva, H. G.; Cristea, I. M. *Anal. Chem.* **2013**, *85* (2), 749–768.
(7) Ngounou Wetie, A. G.; Sokolowska, I.; Woods, A. G.; Roy, U.; Loo, J. A.; Darie, C. C. *Proteomics* **2013**, *13* (3-4), 538–557.
(8) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. *Anal. Chem.* **2014**, *87* (1), 99–118.
(9) Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev.* **2009**, *28* (5), 785–815.
(10) Fitzgerald, M. C.; West, G. M. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1193–1206.
(11) Engen, J. R. *Anal. Chem.* **2009**, *81* (19), 7870–7875.
(12) Konermann, L.; Pan, J.; Liu, Y.-H. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234.
(13) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76* (1), 167–193.
(14) Heck, A. J. R. *Nat. Methods* **2008**, *5* (11), 927–933.
(15) Ben-Nissan, G.; Sharon, M. *Chem. Soc. Rev.* **2011**, *40* (7), 3627–3637.
(16) Snijder, J.; Heck, A. J. R. *Annu. Rev. Anal. Chem.* **2014**, *7* (1), 43–64.
(17) Van Duijn, E. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 971–978.
(18) Vahidi, S.; Stocks, B. B.; Konermann, L. *Anal. Chem.* **2013**, *85* (21), 10471–10478.
(19) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2012**, *85* (1), 2–9.
(20) Schennach, M.; Breuker, K. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 164–168.
(21) Breuker, K.; Brüschweiler, S.; Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, *50* (4), 873–877.
(22) McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. *J. Am. Chem. Soc.* **2015**, *137* (39), 12667-76.
(23) Uetrecht, C.; Rose, R. J.; Duijn, E. van; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39* (5), 1633–1655.
(24) Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.
(25) Jurneczko, E.; Barran, P. E. *The Analyst* **2011**, *136* (1), 20–28.
(26) Niu, S.; Rabuck, J. N.; Ruotolo, B. T. *Curr. Opin. Chem. Biol.* **2013**, *17* (5), 809–817.
(27) Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. *Nat. Methods* **2012**, *9* (11), 1084–1086.
(28) Cubrilovic, D.; Barylyuk, K.; Hofmann, D.; Walczak, M. J.; Gräber, M.; Berg, T.; Wider, G.; Zenobi, R. *Chem. Sci.* **2014**, *5* (7), 2794–2803.

(29) Li, H.; Wongkongkathep, P.; Orden, S. L. V.; Loo, R. R. O.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2060–2068.
(30) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128-9135.
(31) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. *Anal. Chem.* **2014**, *86* (1), 317–320.
(32) Cammarata, M. B.; Brodbelt, J. S. *Chem. Sci.* **2015**, *6* (2), 1324–1333.
(33) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(34) Zhang, Z.; Browne, S. J.; Vachet, R. W. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613.
(35) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(36) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18145–18152.
(37) Zhang, H.; Cui, W.; Gross, M. L. *Int. J. Mass Spectrom.* **2013**, *354–355*, 288–291.
(38) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 343–350.
(39) Zhang, H.; Cui, W.; Gross, M. L.; Blankenship, R. E. *FEBS Lett.* **2013**, *587* (8), 1012–1020.
(40) Blackwell, A. E.; Dodds, E. D.; Bandarian, V.; Wysocki, V. H. *Anal. Chem.* **2011**, *83* (8), 2862–2865.
(41) Zhou, M.; Jones, C. M.; Wysocki, V. H. *Anal. Chem.* **2013**, *85* (17), 8262–8267.
(42) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(43) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(44) Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.; Brodbelt, J. *Anal. Chem.* **2014**, *86* (5), 2534–2542.
(45) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136* (29), 10308–10314.
(46) Warnke, S.; von Helden, G.; Pagel, K. *PROTEOMICS* **2015**, *15* (16), 2804-2812.
(47) Lim, S. M.; Westover, K. D.; Ficarro, S. B.; Harrison, R. A.; Choi, H. G.; Pacold, M. E.; Carrasco, M.; Hunter, J.; Kim, N. D.; Xie, T.; Sim, T.; Jänne, P. A.; Meyerson, M.; Marto, J. A.; Engen, J. R.; Gray, N. S. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 199–204.
(48) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. *Nature* **2013**, *503* (7477), 548–551.
(49) Seeburg, P. H.; Colby, W. W.; Capon, D. J.; Goeddel, D. V.; Levinson, A. D. *Nature* **1984**, *312* (5989), 71–75.
(50) Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmüller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A. *Science* **1997**, *277* (5324), 333–339.

(51) Hammond, D. E.; Mageean, C. J.; Rusilowicz, E. V.; Wickenden, J. A.; Clague, M. J.; Prior, I. A. *J. Proteome Res.* **2015**, *14* (3), 1535–1546.
(52) Ihle, N. T.; Byers, L. A.; Kim, E. S.; Saintigny, P.; Lee, J. J.; Blumenschein, G. R.; Tsao, A.; Liu, S.; Larsen, J. E.; Wang, J.; Diao, L.; Coombes, K. R.; Chen, L.; Zhang, S.; Abdelmelek, M. F.; Tang, X.; Papadimitrakopoulou, V.; Minna, J. D.; Lippman, S. M.; Hong, W. K.; Herbst, R. S.; Wistuba, I. I.; Heymach, J. V.; Powis, G. *J. Natl. Cancer Inst.* **2012**, *104* (3), 228–239.
(53) Chen, C.-C.; Er, T.-K.; Liu, Y.-Y.; Hwang, J.-K.; Barrio, M. J.; Rodrigo, M.; Garcia-Toro, E.; Herreros-Villanueva, M. *PLoS ONE* **2013**, *8* (2), e55793.
(54) Freeman, A. K.; Ritt, D. A.; Morrison, D. K. *Small GTPases* **2013**, *4* (3), 180–185.
(55) Muratcioglu, S.; Chavan, T. S.; Freed, B. C.; Jang, H.; Khavrutskii, L.; Freed, R. N.; Dyba, M. A.; Stefanisko, K.; Tarasov, S. G.; Gursoy, A.; Keskin, O.; Tarasova, N. I.; Gaponenko, V.; Nussinov, R. *Structure* **2015**, *23* (7), 1325–1335.

Vita

Michael B. Cammarata was born in Bethlehem, Pennsylvania on April 14th, 1989 to Don Cammarata and Glenda Boyer. We relocated to Baton Rouge, Lousiana at the age of 5. They remained here for 3 years and then moved to The Woodlands, Texas due to work relocation. Michael graduated from the first class of students from College Park Highschool. After, he attended Trinity University from August 2007 till May 2011 graduating with B.S. in Chemistry. At Trinity Michael took an interest in chemistry research starting in Dr. Michelle Bushey's lab doing capillary electrophoresis research and applying this to monitor phosphorylation of peptides. Additionally, he did research in Dr. Steven Bachrach's lab employing Ab Initio modeling to cucurbitural host-guest systems. Two weeks after graduating from Trinity University Michael started work in Dr. Jennifer Brodbelt's lab beginning his PhD research. Michael studied and developed structural biology methods through the lens of mass spectrometry for his respective PhD research. This interest in structural biology has lead him to Post-Doctoral research in the structural biology field.

Permanent address: 12729 Connemara Lane, Fort Worth, TX, 76244

This dissertation was typed by the author.